

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

Art Unit: 1645

Priority

Please Note: The elected SEQ ID Nos 1, 3 and 5 do not evidence 100% sequence identity with the sequences set forth in the parent application 08/123, 975. Original descriptive support in the parent application, 08/123, 975 was not found for SEQ ID Nos 1, 3 and 5, nor the elected invention of the combination of SEQ ID Nos 1, 3 and 5. The instantly claimed invention is therefore afforded the priority date of May 12, 1999 (provision application) which evidences original descriptive support for each of these sequences.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Db 432 TCAGATGATCAACATCTCTGACATCAATTCGCTGATCTTGGTATCCATCAACCAAA 491
 Qy 492 tcgctgaataactccaataactacatcaacagcgctctgatcgaccagaacgatctc 551
 Db 492 TCGCTGAATTAACATCAAAATCTACATCAACGCGCTCTATCGACCAAGAACGATCTC 551
 Qy 552 caatctgggtaaacatccacgctcttaataacatcatgttcaaacctggagctgtgtgta 611
 Db 552 CAATCTGGGTAAATCAATCAACGCTCTTAATATACATCATGTTCAAACTGGACGGTGTCTGTA 611
 Qy 612 cactcaacgctacatctcggttcaaatctcaatctgttctcgcaaaagaaactggagaaaa 671
 Db 612 CACTCACCGGTACATCTGGATCAAAATCTCAATCTGTTCGACAAAGAACTGAAAGCAAAA 671
 Qy 672 agaatcaaacgactgttagacaacacgctccaaactctgtgtatcccttgaagaactctggg 731
 Db 672 AGAAATCAAGACCTGTACACACACAGTCCAAATCTTGATCTCTGAAAGACTTCTGGGG 731
 Qy 732 tgaactctgcaatagacacaaacgctactacatgtgtgaatctgttagatccgaacaaata 791
 Db 732 TGACTACTGCAGTAGACAAACCGCTACTACATGCTGAATCTGTATCGATCCGAACAAATA 791
 Qy 792 cgttgagctcaacaatgtatgtatcgcggtttacatgttacctgaaggtccggtgttc 851
 Db 792 CGTTGAGCTCAACAAATGTAGGTATCCGCGGTATCATGTACCTGAAAGCTCCGCGTGTTC 851
 Qy 852 tgtatgaccaccaacatctaccgtaactctccctgtaccggtgttagcaaatcatcat 911
 Db 852 TGTATGACTACCAACATCTACTGAACTTCTCCCTGTACCGTGTGTCCAAATTCATCAT 911
 Qy 912 caaagaatacgcgtctgtgttaacaagaacaaatctgttcgaacaatgatcgtgtatacat 971
 Db 912 CAAGAATACGCGTCTGTAAACAGACAAATATGTTGCGAACAAATGATGTGTATCAT 971
 Qy 972 caatgtgtgtgttaagaacaagaatacgcgtctgtgtaccacaaatgtctcgaagctgtgt 1031
 Db 972 CAATGTGTGTGTAAAGAAACAAAGATACCGTCTGTGTACCAATGCTTCTCAGCGTGTGT 1031
 Qy 1032 agaaaagatctgtgtctctgtgaatacccgagctgtgtgaatctgtgtcgaagtaagtgt 1091
 Db 1032 AGAAAGATCTTGTCTCTCTGTGAAATCCCGAGGTTGTGTATGTGTCTCAGGATGTGT 1091
 Qy 1092 aatgaataccaagaagacagaggtatcacttaacaatgtcaaaatgtatctgtcagagcaaa 1151
 Db 1092 AATGAATCAAGAGACAGCAGGTATCACTAAATGCAAAATGATCTGCGAGCAAA 1151
 Qy 1152 caatgtgaacgatacgggttcatcgggttccaccagttcaacaataatcgcgttaactgtgt 1211
 Db 1152 CAATGTGAACGATATCGGTTCATCGGTTCACACAGTTCAACAAATATCGCTAAATGTGT 1211
 Qy 1212 tgcctgaactgtgtacaaatctgcagatcgaacgttctctcgcagctctgtgtgtctgtg 1271
 Db 1212 TGCCTCAACTGTGTACAAATGTCAGATCGAAGCTTCTCTGCGACCTGTGGGTCTCTGTG 1271
 Qy 1272 ggaattatcccggtgtgtatgacggtgttggaagctgcgcgtgtgaa 1317
 Db 1272 GGAGTTATCCCGGTGTGTATGACGGTGTGGGTAACGTCCTGTGAA 1317
 RESULT 3
 AX036243 1330 bp DNA linear PAT 16-NOV-2000
 LOCUS Sequence 22 from Patent EP041149.
 DEFINITION AX036243
 ACCESSION AX036243
 VERSION AX036243.1 GI:11225862
 KEYWORDS
 ORGANISM
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS Kink,J.A., Flitca,J.R., Padhye,N.V., Thalleay,B.S., Stafford,D.C. and Williams,J.A.
 TITLE Vaccine and antitoxin for treatment and prevention of C. Difficile

JOURNAL disease
 Patent: EP 1041149-A 22 OCT-2000;
 OPIDIAN PHARM INC (US)
 FEATURES location/Qualifiers
 source
 1.1330
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1.1317
 /note="Synthetic"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC16478.1"
 /db_xref="GI:11225863"
 /translation="MARLSTPEYIIKNIINISILNRYESNHLIDLSRYASKINIGS
 KVNEDPDKNOIQLFNLESSKIEVILKNAIVNSYENFSTFPIRKYFNISILN
 EYTLINMENSNGKWSLNGELITWLODPEIKORVYFYSOMINSDYINRIVEY
 ITNRLNANSKIYINGRLIDKPISNLSNTHASNNRIFKLDGCDTHRYIWKTFNLPD
 KELNEKEIKLDYDQNSGILKDFWDGIDYDYPYMLNLYDNKYDVANNGVIRGTM
 YLKPGRSVMTNIVLNSLYRGTKEIKKYVAGNDNIVRNDRYINVVYNNKBYR
 LATNQAQAVEKILSLAIEIPDGNLSQVVMKSKNDGITNCKMNLQDNGNDIGFI
 GFHGFNNIAKLIVASNMVNRQIERSSRLTGSWPEIPVDGDMGRPL"
 BASE COUNT 400 a 339 c 246 g 345 t
 ORIGIN
 Query Match 98.6%; Score 1304.4; DB 6; Length 1330;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 12 gatgtcactctcactgaataacatcaagaacatcatcaatcctccatcctgaacctgtg 71
 Db 12 GCTGTACTCTTCACTGATCAATCAAGAACATCATCATCAATCTCAATCTGAACTGCG 71
 Qy 72 ctacgaatccaatcactgtgacgtgtgtcgtctgaagcttccaaatcaaatcgtgttc 131
 Db 72 CTACGAATCCCAATCACTGATCGACCTGTCTCGTACGCTTCCAAATCAACATCGGTTTC 131
 Qy 132 taaagttaactctgcatcgcagatcgcgaagaacatcagatcagctgtgtcaatctgtgaatcttc 191
 Db 132 TAAAGTTAACTTCATCTCGATCGATCGACCAAGAAATCAATCACTGATCTGGAATTTTC 191
 Qy 192 caaatcgaagttaacctgaagaatgtctatcgtatacaactctatgtacgaaaactcttc 251
 Db 192 CAATATGAATTAATCCGAGAAATGATGTATACACTCTATGTATGCAAACTTCTC 251
 Qy 252 caactccttctgtatcgtatccggaatcactcaactcaactcctctctgtgaacatgtgata 311
 Db 252 CACTCTCTTCTGTGATCCGTATCCGAAATACTCAACTCACTCTCTGTGAAATGATA 311
 Qy 312 caccatcaatcaactgtcgtgaagaacaaatctgtgttggaaglatctcgtgaactcgtgtga 371
 Db 312 CACCATCAATCAACTGCAATGGAAGAAATCTGTGTGGAAGATATCTGTGAATCTCGGGTA 371
 Qy 372 aatcatctgacatctgcagagacactcaggaatcaacaacaggtgtgttlatcaaatcactc 431
 Db 372 AATCATCTGCAGCTCTGAGGACACTCAAGGAATCAACAGCGTGTGTATCAATATCTC 431
 Qy 432 tcaagtatcaacaatctcgtactacatcaatcgtgtgatactgttaccatcaccaaaa 491
 Db 432 TCAGATGATCAAACTCTCTGTACATCAATCGCTGTGATCTCTGTACCAATCAACAA 491
 Qy 492 tcgctgaataactccaataactctacatcaacgagcgctctgtatgacagaagaacgatctc 551
 Db 492 TCGCTGAATTAACATCAAAATCTACATCAACGCGCTCTGTATCGACCAAGAACGATCTC 551
 Qy 552 caatctgggtaaacatccacgctcttaataacatcatgttcaaacctggagctgtgtgta 611
 Db 552 CAATCTGGGTAAATCAATCAACGCTCTTAATATACATCATGTTCAAACTGGACGGTGTCTGTA 611
 Qy 612 cactcaacgctacatctcggttcaaatctcaatctgttctcgcaaaagaaactggagaaaa 671
 Db 612 CACTCACCGGTACATCTGGATCAAAATCTCAATCTGTTCGACAAAGAACTGAAAGCAAAA 671

THIS PAGE BLANK (USPTO)

VERSION AM06241.1 GI:11225862

KEYWORDS

SOURCE synthetic construct.

REFERENCE 1 (bases 1 to 1330)

AUTHORS Kink, J. A., Pirca, J. R., Padhye, N. V., Thalley, B. S., Stafford, D. C. and Williams, J. A.

TITLE Vaccines and antitoxin for treatment and prevention of *C. difficile* disease

JOURNAL Patent: EP 1041149-A 22 Oct-2000;

FEATURES OPHIDIAN PHARM INC (US)

Location/Qualifiers

source 1..1330

/organism="synthetic construct"

/db_xref="taxon:32630"

CDS 1..1317

/note="Synthetic"

/codon_start=1

/transl_table=11

/protein_id="CAC16478.1"

/db_xref="GI:11225863"

/translation="MARLLSTPTEYIKNIINTSILNLRYESNHLIDLSHYASKINIGS
KVNFDPIDKNQIOLFNLLESSKIEVLKNAIVYNSMYENFTSFWRIPKYFNSISLNN
EYTIINCMENNSQWKSIVSYGEIILWLODQTEIKQVVFYKSWINISDYINHWIFPT
ITFNRHNNWSTYIPGRILIDQKPIHSLGKHASNNIMFKLDGCRDTHRYIWKYFNLD
KELNKKELKDLVINDOSNSGILKHPWGVYIQDKPYMYLNYDINKYVIWVNVGIRGYM
YLKGPVGSVMYTTNLYLNSSLYKTKF1KKYASGNKUNIVKNDNRVYINVVKNKNEYR
LATNASOAGVEKILSALEIPDVGNLSOVVVMKSKNDGQITNKKCMNLDNNGNDIGFI
GFHQFNNAIKLVASWYNNRIERSRTIGCSWEFIPVDGNGERPL"

BASE COUNT 400 a 339 c 246 g 345 t

ORIGIN

Query Match 98.2%; Score 1302.2; DB 6; Length 1330;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Oy 14  tggcctctaccttcaactgaatcacatcaagaacatcatcaatccctccatcctgaacctgc 73
Db 11  TGCTGTCTACCTTCACTGAATACATCAAGAATCATCAATACCTCCATCTGAACCTGC 70

Oy 74  gctacgaatccaatcacctgatcgacctgtctcgctacgcttccaaaatcaacatcggtt 133
Db 71  GCTACGAATCCAATCACCTGATCGACCTGTCTCGCTACGCTTCCAAAATCAACATCGGT 130

Oy 134  ctaaaagttaacttcgatccgatcgacaagaatcagatccagctgttcaactctggaatctt 193
Db 131  CTAAGTTAACTTCGATCCGATCGACAAGAATCAGATCCAGCTGTTCAAATCTGGAATCTT 190

Oy 194  ccaaaatgaagttaacttcgatccgatcgacaagaatcagatccagctgttcaactctgga 253
Db 191  CCAAAATCGAAGTTATCTCTGAAGATGCTATCGTATACAACCTCTATCTACGAAAACCTT 250

Oy 254  ccacctccttctggatccgatccgaaaacttcaactccatctctctgaacatgaat 313
Db 251  CCACCTCTCTCTGGATCCGATCCGAAAATCTCAACCTCCATCTCTCTGAACATGAAT 310

Oy 314  acaccatcatcaactgcctggaacaaacttctggttggaaagtatctctgaactacgggt 373
Db 311  ACACCATCATCAACTGCATGGAAAACAAATCTGGTGGAAAGATATCTCTGAACATCGGTG 370

Oy 374  aaatcatctggactctgcaggacactcaggaatcaaacagcgtgttgattcaaatact 433
Db 371  AAATCATCTGGACTCTGCAGGACACTCAGGAAATCAACAGCGTGTGTATTCAAATACT 430

Oy 434  ctacagatgaacatctctgactacatcaatcgctggatcttctgttaccatcaccaaca 493
Db 431  CTCAGATGATCAACATCTCTGACTACATCAATCGCTGGATCTCGTACCATCACCAACA 490

Oy 494  atcgtctgaataactccaaaatctacatcaacggccgtctgacgaccagaaacccgatct 553
Db 491  ATCGTCTGAATAACTCCAAAATCTACATCAACGGCCGCTCTGATCGACAGAAACCGATCT 550

Oy 554  ccaatctgggttaactccacgcttctaataacatcatgttcaaacctggacggtgtctgt 613
Db 551  CCAATCTGGGTAACTCCACGCTCTTAATAACATCATGTTCAAACTGGACGGTGTCTGTG 610

Oy 614  acactcaccgctacatctggatcaaatacttcaatctgttgcacaaagaactgaacgaaa 673
Db 611  ACACCTACCCGCTACATCTGGATCAAATACTTCAATCTGTTCGACAAAGAACTGAACGAAA 670

Oy 734  gtagctacctgcagtagcagaaacccgtactacatgctgaatctgttcgacccgaacaa 793
Db 731  GTGACTACCTGCAGTACGACAAACCCGTACTACATGCTGAATCTGTACGATCCGAACAAAT 790

Oy 794  acgttgacgtcaacaatgtaggtatccgggttacatgtacctgaagggtccgcgtggtt 853
Db 791  ACGTTGACGTCAACAATGTAGGTATCCCGGTTACATGTACCTGAAGGTCCCGCTGGTT 850

Oy 854  ctgttatgactaccaacatctacctgaactcttccctgacacgtgtacccaattcatca 913
Db 851  CTGTTATGACTACCAACATCTACCTGAACCTTCCCTGTACCGTGGTACCAATTCATCA 910

Oy 914  tcaagaataacgcgtctggttaacaaggacaatctcgttcgcaacaatgatcgtgtatca 973
Db 911  TCAAGAAATACGCGCTCTGGTAACAAGGACAATATCGTTCGCAACAATGATCGGTATACA 970

Oy 974  tcaatgttgttgttaagaacaaagaataccgtctggctaccgaatgcttccaggtctgtg 1033
Db 971  TCAATGTGTAGTTAAGAACAAGAATACCGTCTGGCTACCAATGCTTCTCAGGCTGGTG 1030

Oy 1034  tagaaagatcttctgtctctggaatcccgagcgttggttaactctgtctcaggtagttg 1093
Db 1031  TAGAAAAGATCTTGTCTGTCTGGAAATCCCGGACGTTGGTAATCTGTCTCAGGTAGTTG 1090

Oy 1094  taatgaatccaagaacgaccagggtatcactaacaagaatgaaatgaatctgcaggaca 1153
Db 1091  TAATGAATCCAAGAACGACCGGTATCACTAACAAGTCAAAATGAATCTGCAGGACA 1150

Oy 1154  acaatggttaacgatctcggttccatcggttccaccagttcaacaatctcgctaaactgg 1213
Db 1151  ACAATGGTAAAGATATCGGTTTCACTCGGTTCCACCAGTTCAACAATATCGCTAAACTGG 1210

Oy 1214  ttgcttcaactggtacaatcgctcagatcgaaacttctctcgcactctgggttgccttt 1273
Db 1211  TTGCTTCAACTGGTACAATCGTCAGATCGAACGTTCTCTCGCACTCTGGGTGCTCTT 1270

Oy 1274  gggagttcatcccggttgatgacgggttggggtgaaacgtccgctgtaa 1320
Db 1271  GGGAGTTCACTCCCGTTGATGACGGTTGGGGTGAACGTCCGCTGTAA 1317
```

THIS PAGE BLANK (USPTO)

RESULT
ARI69140 LOCUS AR169140 1330 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 22 from patent US 6290960.
ACCESSION AR169140
VERSION AR169140.1 G1:17906909
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1330)
AUTHORS Kink, J.A., Thalley, B.S. and Stafford, D.C.
TITLE Vaccine and antitoxin for the treatment of C. difficile disease
JOURNAL Patent: US 6290960-A 22 18-SEP-2001
FEATURES
source Location/Qualifiers
1..1330
/organism="unknown"
BASE COUNT 400 a 339 c 246 g 345 t
ORIGIN

SEQ ID

Query Match 98.2%; Score 1302.2; DB 6; Length 1330;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 tggcctctaccttcaactgaatcacatcaagaacatcatcaatccctccatcctgaacctgc 73
Db 11 TGCTGTCTACCTTCACTGAATACATCAAGAATCATCAATACCTCCATCCTGAACCTGC 70

Qy 74 gctacgaatccaatcaccgatcgacgtgtctcgctacgcttccaaaatcaacatcggtt 133
Db 71 GCTACGAATCCAATCACCTGATCGACCTGTCTCGTACGCTTCCAAAATCAACATCGGTT 130

Qy 134 ctaaaagttaacttcgatccgatcgacaagaatcagatccagctgttcaactctggaatctt 193
Db 131 CTAAAGTTAACTTCGATCCGATCGACAAGAATCAGATCCAGCTGTTCAATCTGGAATCTT 190

Qy 194 ccaaaaatcgaagttatcctgaagaatgctatcgatatcaactctatgtacgaaaaactctt 253
Db 191 CCAAAAATCGAAGTTATCTCTGAAGAATGCTATCGTATACAATCTATGTACGAAAATCTCT 250

Qy 254 ccacctctctcttggaatccgatccggaataacttcaactccatctctctgaacaatgaat 313
Db 251 CCACCTCTCTTCTGGATCCGATCCCGAAATACTTCAACTCCATCTCTGAAACAATGAAT 310

Qy 314 acaccatcatcaactgcatggaaaaaacttctggttggaagatctctgaactacgggtg 373
Db 311 ACACCATCATCAACTGCATGGAAAAAATCTCTGGTTGGAAGATATCTGAACTACGGTG 370

Qy 374 aaatcatctggactctcgaggacactcaggaaatcaaacagcggtgttattcaaaact 433
Db 371 AAATCATCTGGACTCTGAGGACACTCAGGAAATCAAACAGCGGTGTGTTATTCAAATACT 430

Qy 434 ctacgatgatcaacatctctgaactacatcaatcgctggatctctcgtaacatcaacaa 493
Db 431 CTCAGATGATCAACATCTCTGACTACATCAATCGTGGATCTTCGTTACCATCACCAACA 490

Qy 494 atcgcttgaataaactccaaaatctacatcaacgcccgtctgatcgaccagaaacgatct 553
Db 491 ATCGTCTGAATAAATCCAAAATCTACATCAACGGCGCTCTGATCGACCAAGAACCGATCT 550

Qy 554 ccaatctgggttaacatccacgcttctaataacatcatgttcaaaactggacgggtgtcg 613
Db 551 CCAATCTGGGTAACATCCACGCTTCTAATAACATCATGTTCAAAATGGACGGGTGTCTGTG 610

Qy 614 acactcaccgctacatctggatcaaaacttcaatctgttcgacaagaactgaacgaaa 673
Db 611 ACACCTCACCCTACATCTGGATCAAAATACTTCAATCTGTTCGACAAGAACTGAACGAAA 670

Qy 674 aagaaatcaagacactgtacgacaaccagtcgaattctggtatctcgaagaactctcg 733
Db 671 AAGAAATCAAGACCTGTACGACAACAGTCCAATCTGGTATCTCGAAGACTTCTCGGG 730

Qy 734 gtgactacctgcagtagcacaacacggtactacatgctgaatctgtacgacgcgaacaaat 793
Db 731 GTGACTACCTGCAGTACGACAACCGGTACTACATGCTGAATCTGTACGATCCGAACAAAT 790

Qy 794 acggttagctcaacatgtaggatccgcggttacatgtacctgaagggtccgctgggtt 853
Db 791 ACGTTGACGTCAACAATGTAGGTATCCGCGGTACATGTACCTGAAGGTCGCGTGGTT 850

Qy 854 ctgttatgactaccaacatctacctgaactcttccctgtacggtggtacaaaattcata 913
Db 851 CTGTTATGACTACCAACATCTACCTGAACCTTCCCTGTACCGTGGTACCAAAATTCATCA 910

Qy 914 tcaagaatacgcgtctggttaacaaggacaatatcggttcgcaacaatgatcggtatata 973
Db 911 TCAAGAAATACGCGTCTGGTAACAAGGACAATATCGTTCGCAACAATGATCGGTATATA 970

Qy 974 tcaatgtttagttaagaacaaagaataaccgctctggctaccaatgcttctcaggctgg 1033
Db 971 TCAATGTTGTAGTTAAGAACAAGAAATACCGTCTGGCTACCAATGCTTCTCAGGCTGGT 1030

Qy 1034 tagaaaagatcttgtctgctctggaatcccggaactgtgtaactctgctcaggtagttg 1093
Db 1031 TAGAAAAGATCTTGTCTGCTCTGGAATCCCGGACGTTGGTAATCTGTCTCAGGTAGTTG 1090

Qy 1094 laaagaaatccaagaacgaccaggtatcactaacaatacgaataatgaatctgcaggaca 1153
Db 1091 TAATGAAATCCAAGAACGACGAGGTATCACTAACAAATGCAAAATGAATCTGCAGGACA 1150

Qy 1154 acaatggtaacgatatacggtttcactcggtttccaccagttcaacaatatcgctaaactg 1213
Db 1151 ACAATGGTAACGATATCGGTTTCATCGGTTTCCACCAAGTCAACAATATCGCTAAACTGG 1210

Qy 1214 ttgcttccaactggtaacatcgcagatcgaaacttccctctcgacactctgggttgcctt 1273
Db 1211 TTGCTTCCAACCTGGTACAATCGTCAGATCGAACGTTCCCTCTCGCACTCTGGGTGCTCTT 1270

Qy 1274 gggagttcatcccggttgatgacggttgggggtgaacgtccgctgtaa 1320
Db 1271 GGGAGTTCA'PCCCGGTTGATGACGGTGGGGTGAACGTCCGCTGTAA 1317

THIS PAGE BLANK (USPTO)

RESULT 4
 XXU22962
 LOCUS
 DEFINITION Synthetic botulinum neurotoxin serotype A Hc fragment (botA) gene.
 ACCESSION U22962
 VERSION U22962.1
 KEYWORDS GI:733428
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 1338)
 AUTHORS Clayton, M.A., Clayton, J.M., Brown, D.R. and Middlebrook, J.L.
 TITLE Protective vaccination with a recombinant fragment of Clostridium botulinum neurotoxin serotype A expressed from a synthetic gene in Escherichia coli
 JOURNAL Infect. Immun. 63 (7), 2738-2742 (1995)
 MEDLINE 95310035
 REFERENCE 2 (bases 1 to 1338)
 AUTHORS Brown, D.R.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-1995) Douglas R. Brown, Toxinology, U. S. Army Medical Research Institute of Infectious Diseases, Bldg. 1425, Fort Detrick, Frederick, MD 21702-5011, USA
 FEATURES
 source
 1. 1338
 /organism="synthetic construct"
 /db_xref="taxon:32630"

SEQID1

1995

Query Match 98.4%; Score 1311; DB 12; Length 1338;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 16 cgtctgctgtctaccttctactgaatacatcaagaacatcatcaatctccatctggaac 75
 Db 15 CGTCTGCTGTCTACCTTCACTGAATACATCAAGAATCATCAATACCTCCATCTGAAC 74
 Oy 76 ctgcgctacgaatccaatcactctgacgactgctcgtcgaagcttccaaatcaaacatc 135
 Db 75 CTGCGCTACGAATCCAATCACCTGATCGACCTGTCTCGCTACGCTTCCAAATCAACATC 134
 Oy 136 ggttctaaagttaacttcgacgacgacgacgacgacgacgacgacgacgacgacgacgac 195
 Db 135 GGTTCCTAAAGTTAACTTCGATCCGATCGACAAGAAATCAGATCCAGCTGTCTCAATCTGAA 194
 Oy 196 tcttccaaatcgaagttaacttcgacgacgacgacgacgacgacgacgacgacgacgacgac 255
 Db 195 TCTTCCAAATCGAAGTTATCTCGAAGAAATGCTATCGTATCAACTCTATGACGAAATC 254
 Oy 256 tcttccaaatcgttctgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 315
 Db 255 TCTTCCAAATCGAAGTTATCTCGAAGAAATGCTATCGTATCAACTCTATGACGAAATC 314
 Oy 316 gaatacaccatcatcaactgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 375
 Db 315 GAATACACCATCATCAACTGATGAAACAAATCTGTTGGAAGATCTCTGAACTAC 374
 Oy 376 ggtgaaatcatctgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 435
 Db 375 GGTGAAATCATCTGGACTCTCGAGGACACTCAGGAAATCAACAGCGTGTGTATTCAA 434
 Oy 436 tactctcagatgatcaacatctctgactacatcaactcgtggaactctctgttaccatcacc 495
 Db 435 TACTCTCAGATGATCAACATCTCTGACTACATCAATCGCTGGATCTCTGTTACCATCACC 494
 Oy 496 aacaatcgtctgaataaactccaaatctacatcaacgacgacgacgacgacgacgacgacgac 555
 Db 495 AACATCGTCTGAATAACTCCAAATCTACATCAACGCGCTGTGATCGACGAGAACCG 554
 Oy 556 atctccaaatctgggttaacatccacgcttcttaataacatcatgttcaaaactggaggtgt 615
 Db 555 ATCTCCAATCTGGGTAACTCCACGCTTCTAATAACATCATGTCAAACTGGACGGTGT 614
 Oy 616 cgtgacactcaccgctacatctggtatcaaaacttcaactctgttgcgaagaagaactgaac 675
 Db 615 CGTGACACTCACCGCTACATCTGGATCAAACTTCAATCTGTTGACAAAGAACTGAAC 674
 Oy 676 gaaaaagaaatcaaaagacctgtacgacaaacagctccaattctggtatctctgaaagacttc 735
 Db 675 GAAAAAGAAATCAAGACCTGTACGACAAACAGTCCAATCTGGTATCTGTAAGACCTTC 734
 Oy 736 tgggggtgactacctcgtacgacaaacgctactacatgctgaatctgacgacgacgacgac 795
 Db 735 TGGGGTGACTACCTGCAGTACGACAAACCGTACTACATGCTGAATCTGTACGATCCGAAC 794
 Oy 796 aaatacgttgacgctcaaatgttaggtatccgctgtacatgtacgtgaaaggtccgct 855
 Db 795 AAATACGTTGACGCTCAAAATGTAGGTATCCGCGTTACATGTACCTGAAAGGTCCGCT 854
 Oy 856 ggttctgttatgactaacaacatctactgaactcttccctgtaccgtggtaccaaattc 915
 Db 855 GGTTCGTGTATGACTACCAACATCTACCTGAATCTTCCCTGTACCGTGTACCAATTC 914
 Oy 916 atcatcaagaatcgcgtctggttaacaaaggaataatcgttcgcaacatgatcgtgta 975
 Db 915 ATCATCAAGAAATACGCGTCTGGTAACAGGACAATATCGTTCGCAACATGATCGTGTA 974
 Oy 976 tacatcaatgtttagtttaagaacaaagaataccgctgtggttccaatgtcttcaggct 1035
 Db 975 TACATCAATCTGTAGTTAAGAACAAAGATACCGTCTGGCTACCAATGCTTCTCAGGCT 1034
 Oy 1036 ggtgtagaagaatcttctgtctctggaatcccggaactgtggttaactctgtctcaggta 1095
 Db 1035 GGTGTAGAAGATCTTGTCTGCTTGAATCCCGGACGTGGTAAATCTGTCTCAGGTA 1094
 Oy 1096 gttgtaataaatacgaagaacgacgggtatcactaacaataatgcaaaatgaatctgcag 1155
 Db 1095 GTTGTAAATGAATCCAAAGACGACGAGGTATCACTAACAAATGCAAAATGAATCTGCAG 1154
 Oy 1156 gacaacaatggttaacgatatcgttttcaatcgggtttccaccagttcaacaatatcgctaaa 1215
 Db 1155 GACAACAATGTTAAGGATATCGGTTTCATCGGTTTCCACCAATTCACCAATATCGCTAAA 1214
 Oy 1216 ctggttgcgttccaaactggtacatcgtcagatcgaacgttctctcgaactctgggttgc 1275
 Db 1215

Clayton et al

THIS PAGE BLANK (USPTO)

RESULT 4
 XXU22962
 LOCUS XY 62 1338 bp DNA line: SYN 08-NOV-1995
 DEFINITION Synthetic botulinum neurotoxin serotype A Hc fragment (botA) gene, complete cds.
 ACCESSION U22962
 VERSION U22962.1 GI:733428
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 1338)
 AUTHORS Clayton, M.A., Clayton, J.M., Brown, D.R. and Middlebrook, J.L.
 TITLE Protective vaccination with a recombinant fragment of Clostridium botulinum neurotoxin serotype A expressed from a synthetic gene in Escherichia coli
 JOURNAL Infect. Immun. 63 (7), 2738-2742 (1995)
 MEDLINE 95310035
 REFERENCE 2 (bases 1 to 1338)
 AUTHORS Brown, D.R.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-1995) Douglas R. Brown, Toxinology, U. S. Army Medical Research Institute of Infectious Diseases, Bldg. 1425, Fort Detrick, Frederick, MD 21702-5011, USA
 FEATURES
 source
 1. 1338
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /gene="botA"
 /codon_start=1
 /transl_table=11
 /product="botulinum neurotoxin serotype A Hc fragment"
 /protein_id="AA080610.1"
 /db_xref="GI:733429"
 /translation="MARLLSTFTEYIKNIINTSILNLRYESNHLIDLSRYASKINIGSKVNFDPIDKNQIQFLNLESSKIEVILKNAIVYNSMYENFSTFSWIRIPKYFNSISLNN EYTIINCMEENSGWKVSLNYGEIITLQDTQEIQRVVFYKYSQMINISDYINRWIFVT ITNRLNNSKIYINGRLIDQKPIISNLGNIHASNNIMFKLDGCRDTHRYIWKYFNLFED KELNEKEIKDLYDNQSNISGILKDFWGDYLYQDKPYMLNLYDPNKYVDVNNVVGIRGYM YLKGPRGSMVTNNIYLNSSLYRGTFKIIKKYASGNKDNIVRNNDRYVINNVKNEYR LFNATASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNCKCKMNLQDNNNGNDIGFI GFHQFNIAKLVASNWNQRQIERSRSLTLCGSWEFIPVDDGWGERPL"
 BASE COUNT 401 a 342 c 249 g 346 t
 ORIGIN

Query Match 98.2%; Score 1302.2; DB 12; Length 1338;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 tggcctctaccttcactgaatacatcaagaacatcatcaatacctccatcctgaacctgc 73
 || |||||
 Db 19 TGCTGTCTACCTTCACTGAATACATCAAGAACATCATCAATACCTCCATCCTGAACCTGC 78
 Qy 74 gctacgaatccaatcacctgatcgacctgtctcgctacgcttccaaaatcaaatcgggt 133
 |||||
 Db 79 GCTACGAATCCAATCACCTGATCGACCTGTCTCGCTACGCTTCCAAAATCAACATCGGTT 138
 Qy 134 ctaaaagttaacttcgatccgatcgacaagaatcagatccagctgttcaatctggaatctt 193
 |||||
 Db 139 CTAAAGTTAACTTCGATCCGATCGACAAGAATCAGATCCAGCTGTTCAATCTGGAATCTT 198
 Qy 194 ccaaaatcgaagtattcctgaagaatgctatcgatatacaactctatgtacgaaaacttct 253
 |||||
 Db 199 CCAAAATCGAAGTTATCCTGAAGAATGCTATCGTATACAACCTCTATGTACGAAAACCTCT 258
 Qy 254 ccacctctcttctggatccgatcccgaaataacttcaactccatctctctgaacaatgaat 313
 |||||
 Db 259 CCACCTCCTTCTGGATCCGTATCCCGAAATACTTCAACTCCATCTCTCTGAACAATGAAT 318
 Qy 314 acaccatcatcaactgcgatggaacaaatcttggttggaagtattctctgaactacgggtg 373
 |||||
 Db 319 ACACCATCATCAACTGCATGGAAAACAATTCTGGTTGGAAGTATCTCTGAACCTACGGTG 378
 Qy 374 aaatcatctggactctgcaggacactcaggaaatcaaacagcgtgtgtattcaataact 433
 |||||
 Db 379 AAATCATCTGGACTCTGCAGGACACTCAGGAAATCAAACAGCGTGTGTATTCAAATACT 438
 Qy 434 ctcatgatcatcaatctctgactacatcaatcgctggatcttctgttaccatcaccaaca 493
 |||||
 Db 439 CTCAGATGATCAACATCTCTGACTACATCAATCGCTGGATCTTCGTTACCATCACCAACA 498
 Qy 494 atcgtctgaataactccaaatctacatcaacggccgctctgatcgaccagaacccgatct 553
 |||||
 Db 499 ATCGTCTGAATAACTCCAAATCTACATCAACGGCCGCTGTGATCGACCAGAAACCGATCT 558
 Qy 554 ccaatctgggttaacatccacgcttcttaataacatcatgttcaaaactggacggtgtcggtg 613
 |||||
 Db 559 CCAATCTGGGTAAACATCCACGCTTCTAATAACATCATGTTCAAACTGGACGGTTGTCTGTG 618
 Qy 614 aactcaccgctacatctggatcaaatcttcaatctgttcgacaagaactgaacgaaa 673
 |||||
 Db 619 AACTCACCCTACATCTGGATCAAACTTCAATCTGTTCGACAAGAACTGAACGAAA 678
 Qy 674 aagaaatcaagacctgtacgacaaccagtcgaattctggtatcctgaaagacttctggg 733

Clayton et al
 1995

SEQIDS

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 17:10:08 : Search time 7254.67 Seconds

(without alignments)
3842.235 Million cell updates/sec

Title: US-09-611-419A-1

Perfect score: 1332

Sequence: 1 gaatcgaacagctgcgtc.....aacgtccgcgtgaagattc 1332

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sy:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_or:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_pl:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	1311	98.4	1330	6	AR000029	AR000029 Sequence
2	1311	98.4	1330	6	AR169140	AR169140 Sequence
3	1311	98.4	1330	6	AX036243	AX036243 Sequence
4	1311	98.4	1338	12	XX022962	U22962 Synthetic b
5	1311	98.4	1402	6	AR000030	AR000030 Sequence
6	1311	98.4	1402	6	AR169141	AR169141 Sequence
7	1311	98.4	1402	6	AX036246	AX036246 Sequence
8	727.6	54.6	4292	1	CBNT0X	X52066 Clostridium
9	727.6	54.6	4835	1	CLONEUR	M30196 C.botulinum
10	726.4	54.5	3891	6	AR000031	AR000031 Sequence
11	726.4	54.5	3891	6	AR169142	AR169142 Sequence
12	726.4	54.5	3891	6	AX036248	AX036248 Sequence
13	662	49.7	4067	1	CBNT0X	X73423 C.botulinum
14	423.2	31.7	1313	6	AF251281	AF251281 Synthetic
15	421.6	31.7	1313	6	A58946	A58946 Sequence 6
16	227.8	17.1	4073	1	CBNT0X	X68262 C.barrati ge
17	227.2	17.1	1293	6	A58945	A58945 Sequence 5
18	227.2	17.1	4199	1	CLOBONT	L35496 Clostridium
19	227.2	17.1	4209	1	CBNT0X	X81714 C.botulinum
20	195	14.6	3835	1	AB037704	AB037704 Clostridi
21	195	14.6	3835	1	AB037705	AB037705 Clostridi
22	195	14.6	3835	1	AB037706	AB037706 Clostridi
23	195	14.6	3835	1	AB037707	AB037707 Clostridi
24	195	14.6	3835	1	AB037708	AB037708 Clostridi
25	195	14.6	3835	1	AB037709	AB037709 Clostridi
26	195	14.6	3835	1	AB037710	AB037710 Clostridi
27	195	14.6	3835	1	AB037711	AB037711 Clostridi
28	195	14.6	3835	1	AB037712	AB037712 Clostridi
29	195	14.6	3835	1	AB037713	AB037713 Clostridi
30	195	14.6	3835	1	AB037714	AB037714 Clostridi
31	195	14.6	4030	1	CBNT0X	X62083 C.botulinum
32	193.4	14.5	4017	6	AX088262	AX088262 Sequence
33	193.4	14.5	4017	6	CBNT0X	X62083 C.botulinum
34	190.2	14.3	3949	1	AB039264	AB039264 Clostridi
35	190.2	14.3	3949	1	CBNT0X	X62083 C.botulinum
36	180.8	13.6	4234	1	CLONEUR	X62083 C.botulinum
37	180.8	13.6	9335	1	CBY13631	Y13631 Clostridium
38	166.4	12.5	1359	6	I28431	I28431 Sequence 3
39	166.4	12.5	3712	6	A49987	A49987 Sequence 4
40	166.4	12.5	3754	6	A37074	A37074 Sequence 17
41	166.4	12.5	3754	6	A42478	A42478 Sequence 6
42	166.4	12.5	3769	6	A37075	A37075 Sequence 18
43	166.4	12.5	3769	6	A49988	A49988 Sequence 5
44	166.4	12.5	4366	6	A42484	A42484 Sequence 12
45	166.4	12.5	4378	6	A42481	A42481 Sequence 9

ALIGNMENTS

RESULT 1

AR000029 AR000029 1330 bp DNA 1linear PAT 04-DEC-1998

LOCUS Sequence 22 from patent US 5736139.

DEFINITION AR000029

ACCESSION AR000029

VERSION AR000029.1 GI:3962560

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1. (bases 1 to 1330)

AUTHORS Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.

TITLE Treatment of Clostridium difficile induced disease

JOURNAL Patent: US 5736139-A 22 07-APR-1996;

FEATURES

source 1. 1330

Location/Qualifiers

1. 1330

BASE COUNT 400 a 339 c 246 g 345 t

ORIGIN

Query Match 98.4%; Score 1311; DB 6; Length 1330;

Best local Similarity 100.0%; Pred. No. 0;

Matches 1311:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
QY	16	cgctcgtctctac	ctcaatgatacatcaagaacatcaatcaatccatccatcc	75				
Db	7	CGCTCGTCTAC	CTTCACTGATATCATCAAGACATCATCAATCACTCATCTCAAC	66				
QY	76	ctgcgtacgaatca	atcaatcaatcgtacgtctctcgtacgcttccaaatcaatc	135				
Db	67	CTCGGTACGAAT	CAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	126				
QY	136	ggttcaaaagtta	ctcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	195				
Db	127	GGTTCTAAATTA	CTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	186				
QY	196	tcttccaaatcga	agttatccctgaagatgctatcgatatacaatcgtatgaagaa	255				
Db	187	TCTTCCAAAT	TGAAGTATCTGATATCAATCAATCAATCAATCAATCAATCAAT	246				
QY	256	tcttccaaatcgt	ctcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	315				
Db	247	TTCCTCACCC	CTTCTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT	306				
QY	316	gaatacacca	atcaatcgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	375				
Db	307	GAATACACCA	TCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	366				
QY	376	ggtgaata	atcctcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	435				
Db	367	GGTGAATAT	CTGATCTGAGACACTGAGACACTGAGACACTGAGACACTGAGAC	426				
QY	436	tacttcaag	tgtacatctctcgtacgtacgtacgtacgtacgtacgtacgtacgt	495				
Db	427	TACTCTAGAT	GTATCAATCTGATCAATCTGATCAATCTGATCAATCTGATCAAT	486				
QY	496	aaacaatcgt	ctcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat	555				
Db	487	AACAATCTGT	GAATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	546				
QY	556	atctcaatc	gtgtacatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat	615				
Db	547	ATCTCAATCT	GTGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	606				
QY	616	cgtaaac	atcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	675				
Db	607	CGTGAAC	CTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	666				
QY	676	gaaaga	aaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat	735				
Db	667	GAAAAA	GAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	726				
QY	736	tgggtg	taactcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	795				
Db	727	TGGGTG	TGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	786				
QY	796	aaatac	gtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	855				
Db	787	AAATAC	GTGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	846				
QY	856	ggttctg	ttaacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	915				
Db	847	GGTTCTGT	TGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	906				
QY	916	atcata	gaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat	975				
Db	907	ATCATCA	GAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	966				
QY	976	tacata	atgttgaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat	1035				
Db	967	TACATCA	ATGTGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	1026				
QY	1036	ggttga	aaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat	1095				
Db	1027	GGGTGA	AAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	1086				

QY	1096	gttgaatga	aatccaaagcagccagggtatcaataaalycaaatgaaatcgtcag	1155				
Db	1087	GTGTATGA	AATCCAAAGACGACGATATCACTAAACAAATGAAATGATCTCAG	1146				
QY	1156	gacaaatg	gttaacgtatcgttttcaatcgttttcaacagttcaacataltcgtcaaa	1215				
Db	1147	GACAAATG	TAAAGATATCGTTTCAATCGTTTCAATCGTTTCAATCGTTTCAAT	1206				
QY	1216	ctggtt	cttccaaatcgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	1275				
Db	1207	CTGGTCTT	CAACTGATCAATCGATCAATCGATCAATCGATCAATCGATCAATCG	1266				
QY	1276	tcttgg	agttcatcccggttgaacggttgggtgaacgttcgtcgtgttaa	1326				
Db	1267	TCTTGGAG	TTCAATCCCGTTGATGACGTTGGGTGAACGTCGCTGTAA	1317				
RESULT 2								
AR169140	LOCUS	AR169140	Sequence 22 from patent US 6290960.	1330 bp	DNA	linear	PAT 17-DEC-2001	
AR169140	DEFINITION	Sequence 22 from patent US 6290960.						
AR169140	ACCESSION	AR169140						
AR169140.1	VERSION	AR169140.1	GI:17906909					
Unknown.	KEYWORDS	Unknown.						
Unknown.	SOURCE	Unknown.						
Unclassified.	ORGANISM	Unclassified.						
1 (bases 1 to 1330)	REFERENCE	1 (bases 1 to 1330)						
Kink, J.A., Thalley, B.S. and Stafford, D.C.	AUTHORS	Kink, J.A., Thalley, B.S. and Stafford, D.C.						
Vaccine and antitoxin for the treatment of C. difficile disease	TITLE	Vaccine and antitoxin for the treatment of C. difficile disease						
Patent: US 6290960-A 22 18-SEP-2001.	JOURNAL	Patent: US 6290960-A 22 18-SEP-2001.						
Location/Qualifiers	FEATURES	Location/Qualifiers						
1..1330	source	1..1330						
BASE COUNT	400 a	339 c	246 g	345 t				
ORIGIN								
Query Match	98.4%:	Score 1311:	DB 6:	Length 1330:				
Best Local Similarity	100.0%:	Pred. No. 0:						
Matches 1311:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
QY	16	cgctcgtctctac	ctcaatgatacatcaagaacatcaatcaatccatccatcc	75				
Db	7	CGCTCGTCTAC	CTTCACTGATATCATCAAGACATCATCAATCACTCATCTCAAC	66				
QY	76	ctgcgtacgaatca	atcaatcaatcgtacgtctctcgtacgcttccaaatcaatc	135				
Db	67	CTCGGTACGAAT	CAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	126				
QY	136	ggttcaaaagtta	ctcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	195				
Db	127	GGTTCTAAATTA	CTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	186				
QY	196	tcttccaaatcga	agttatccctgaagatgctatcgatatacaatcgtatgaagaa	255				
Db	187	TCTTCCAAAT	TGAAGTATCTGATATCAATCAATCAATCAATCAATCAATCAAT	246				
QY	256	tcttccaaatcgt	ctcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	315				
Db	247	TTCCTCACCC	CTTCTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT	306				
QY	316	gaatacacca	atcaatcgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	375				
Db	307	GAATACACCA	TCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	366				
QY	376	ggtgaata	atcctcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgt	435				
Db	367	GGTGAATAT	CTGATCTGAGACACTGAGACACTGAGACACTGAGACACTGAGAC	426				
QY	436	tacttcaag	tgtacatctctcgtacgtacgtacgtacgtacgtacgtacgtacgt	495				

Db 427 TACTCTAGATGATCAACATCTCTGACATCATCATCGTGTGATCTTGTACCATCACC 486
Oy 496 aacatcgtctgataactcaaatctacatacaacgagcgtctgatacgaacgaacccg 555
Db 487 AACAAATGCTGTGAATTAATCTCAAAATCTACATACAGCGCGTGTGATCGACCGAAACCG 546
Oy 556 atctcaatctgggtaacatccacgctctcaataacatcgttcaactgtaacgagttgt 615
Db 547 ATCTCCAAATCTGGGTAAATCCACGCTCTGTATTAACATCATGTTTCAAACTGAGAGGTTGT 606
Oy 616 cgtgaacatccacgctacatctgatacaaatctcaatctgtctgagaagaacgtaac 675
Db 607 CGGACACTCACCGCTACATCTGGATCAATCTTCAATCTGTCGACAAAGAACTGTAAC 666
Oy 676 gaaagaacatcaaaagaccgtgatacgaacacagctccaatctgtgtatccctgaagaacttc 735
Db 667 GAAAAAGAAATCAAAACCTGTACGACACCACTTCAATCTGTGTATCTCGAAGACTTC 726
Oy 736 tggggtgactacgtgatacgaacacgctacatacgtgatacgtgatacgtgatacgtgatac 795
Db 727 TGGGGTACTACTGACGATGACGAAACCGTACTACATGCTGATGCTGATGCTGATGCTGATG 786
Oy 796 aatcactgtaagcgtcaacaaatgtagtatacgcggttaacatgtaacgtgaaggtccggt 855
Db 787 AATATGCTGACGCTCAACATCTAGGTATCCGCGGTACATGTAACCGAAGGTCGCGGT 846
Oy 856 ggtctgatacgtacacacatctacacacacacacacacacacacacacacacacacacacac 915
Db 847 GGTCTGTATGACTACCAACATCTACCTGTAACCTGTAACCTGTAACCTGTAACCTGTAACCTG 906
Oy 916 atcaataagaatacgcgtctgatacgaagaacatacgtctgagaacatacgtctgatacgtgata 975
Db 907 ATCATCAAGAAATACGCGTCTGTGTAAACGAGAAATATCTGTGCAACATGATGATCGTGA 966
Oy 976 tacatacgtctgatacgaagaacatacgcgtctgatacgaagaacatacgtctgatacgtgata 1035
Db 967 TATATCAATGTTGTATGTAAGAAACAAAGATATCGTCTGCTCAATGCTTCTCAAGCT 1026
Oy 1036 ggtgataagaatacgtctgatacgaagaacatacgcgtctgatacgaagaacatacgtctgata 1095
Db 1027 GGTGTGAAAGATCTTGTCTGTCTGTGAAATCCGAGCTGTGTATCTGTCTCAAGCT 1086
Oy 1096 gttgataagaatacgaagaacatacgcgtctgatacgaagaacatacgtctgatacgtgatacgtgata 1155
Db 1087 GTTGTATGTAATCAACGAGAAACGAGGATCTACATACAAATGCAAAATGATGATCTGACG 1146
Oy 1156 gatacaatgtaacgatacgtctgatacgaagaacatacgcgtctgatacgaagaacatacgtgata 1215
Db 1147 GACAAACATGTAAGATATCGGTTTCAATCGGTTTCAACGAGTTCACCAATATGCTGATAA 1206
Oy 1216 ctgtgctctcaacgtgatacgaagaacatacgcgtctgatacgaagaacatacgtctgatacgtgata 1275
Db 1207 CTGTGTCTCTCAACGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1266
Oy 1276 tcttggagttcaccggtctgatacgaagaacatacgcgtctgatacgaagaacatacgtctgata 1326
Db 1267 TCTTGGGAGTCTATCCCGGTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1317

RESULT 3
AX036243 1330 bp DNA 1linear PAT 16-NOV-2000
LOCUS AX036243 Sequence 22 from Patent EP1041149.
DEFINITION AX036243
ACCESSION AX036243.1 GI:11225862
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

synthetic construct.
synthetic construct.
artificial sequence.
1 (bases 1 to 1330)
Kink,J.A., Firta,J.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and
Williams,J.A.
Vaccine and antitoxin for treatment and prevention of c. Difficile

JOURNAL
Patent: EP 1041149-A 22 04-OCT-2000;
OPHIDIAN PHARM INC (US)
FEATURES
Location/Qualifiers
Source
1..1330
/organism="synthetic construct"
/db_xref="taxon:32630"
1..1317
/note="Synthetic"
/codon_start=1
/transl_table=11
/protein_id="CAC16478.1"
/db_xref="GI:11225863"
/translation="MARLSTFEYIKNIITNSILNRYESNHLIDLSRYASKINIGS
KYNFIDIKNOIOLFESSKIEVLKNIIVNMYENSTSPRIIPKPNISLNN
EYTIINCMENNSMKRSLNGETIPILOPDEKOVREKYSOMINISPIYINRIVT
ITNNRINSKITYINGRLIDOKPLSNLGNHASNNIMFKLDGCDHRTYIWKTFINLD
KELNEKIRLDYDNOSSNGLKDFWGDVLYQDPYMLNLYDPNKVVDVNVNGIRGYM
YKGPSPGYWTTIYLNSSLYRGTFTIIRKYSAGNDNIVRNDRVYIVNVVSKYR
LATNASOAGVEKLLSALIEIPDVGNLSQVVVMSKNDQGITNCKMKMLDNNNDIGFI
GFHQFNINIKIVASNNYNRQIERSRSLTLCDSWEFTIVDDGMBRPL"

BASE COUNT 400 a 339 c 246 g 345 t
ORIGIN

Query Match 98.4%; Score 1311; DB 6; Length 1330;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 16 cgtctgctgctacacacacacacacacacacacacacacacacacacacacacacacacacacacac 75
Db 7 CGTCTGCTGTACCTACCTTACCTGATATCATCATCAAGAAATCATCATCATCTCATCTGTAAC 66
Oy 76 ctgcgtcgaatccacacacacacacacacacacacacacacacacacacacacacacacacacacac 135
Db 67 CTGGGCTACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATC 126
Oy 136 gttctaaagttaactcgcgtctgatacgaagaacatacgcgtctgatacgaagaacatacgcgtctgata 195
Db 127 GTTCTTAAATGTAATCTGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 186
Oy 196 tcttccaaatcgaagttatctcctgagaagatgtaacgtatatacaactatgatacgaagaac 255
Db 187 TCTTCCAAATTCGAAGTTATTCGTAAGATGCTATGTAACAACTGATATGTAACAAAC 246
Oy 256 ttctcaactcctctgatacgcgtctgatacgaagaacatacgcgtctgatacgaagaacatacgcgtctgata 315
Db 247 TTTCCACCTCTCTGTGATCCGATCCGAAATTAATCAATCTCATCTCTGTAACAAAT 306
Oy 316 gaataacacatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatca 375
Db 307 GAATACACCATCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 366
Oy 376 gttgaaatcctgagactctgacgagacacacacacacacacacacacacacacacacacacacacac 435
Db 367 GGTGAATATATTCGATCTGACGAGACATCAGGAATCAACGCGTGTGATTTGATAA 426
Oy 436 tacttcaagatgatacactctgatacgaagaacatacgcgtctgatacgaagaacatacgcgtctgata 495
Db 427 TACTCTCAGATGATCAACATCTGTGATCATCATCATCATCATCATCATCATCATCATCATCATCAT 486
Oy 496 aacaaatgctgataaactcaacaaatctacatacaacgagcgtctgatacgaagaacacg 555
Db 487 AACAAATGCTGTGAATTAATCTCAAAATCTACATACAGCGCGTGTGATCGACCGAAACCG 546
Oy 556 atctcaatctgggtaacatccacgctctcaataacatcgttcaactgtaacgagttgt 615
Db 547 ATCTCCAAATCTGGGTAAATCCACGCTCTGTATTAACATCATGTTTCAAACTGAGAGGTTGT 606
Oy 616 cgtgaacatccacgctacatctgatacaaatctcaatctgtctgagaagaacgtaac 675
Db 607 CGGACACTCACCGCTACATCTGGATCAATCTTCAATCTGTCGACAAAGAACTGTAAC 666

```

QY 676 gaaaaaagaatcaagaacctgtacgacacacgaagtcacattctgtgtatctcgtgaagacttc 735
    |||
Db 667 GAAAGAAAGAAATCAAGACCTGTACGACACACGATCCCAATTCGTGTATCTGAAAGACTTC 726
QY 736 tgggggtgacctgacgtacgacaaacgtaactacatgctgaatctgctcgatccgaac 795
    |||
Db 727 TGGGGTACTACTGCTGCTACGACAAACCGTACTACTGCTGAATCTGTACGATCCGAAAC 786
QY 796 aaatacgtctgacgtcaacaaatgtagtataccggtgtacatgtaactgtaagagtcgcgct 855
    |||
Db 787 AAATACGTGACGTCACCAATGTAGTATCCGGGGTATACATGTACTGAAAGTCCGCGT 846
QY 856 ggtctgtatgtacctacaacatctaccgttaactctccctgtacgctgtgtaccaatc 915
    |||
Db 847 GGTTCGTGTATGACTACCAACATCTTACCTTAACCTTCCTGACCGTGTACCAAAATTC 906
QY 916 ataatcaagaatacagcgtctgtgttaacaagaacataatgctgtgcaacaatgtaactgta 975
    |||
Db 907 ATCATCAAGAAATACGCGCTGTGTAACAAGACAAATATCTGTCGCAACATGATCGTGA 966
QY 976 tacatcaatgtctgtagttaagaacaaataccgtctgtgtaacaaatgcttctcaagct 1035
    |||
Db 967 TACATCAATGTGTAGTATGAAGAAACAAAGATACCGTCTGCTACCAATGCTTTCAGGCT 1026
QY 1036 ggtgtgtaagaagatctgtctgtctgtgaatacccggaagcttggtaactgtctcgaagta 1095
    |||
Db 1027 GGTGTGAAAGAAATCTGTGTCTGTGAAATCCCGGACGTGTGTAATCTGTCTCAGGTA 1086
QY 1096 gtgtataatgaatccaagaacgacacgaaggtatcactacaatgtaagaatctgcaag 1155
    |||
Db 1087 GTTGTATGAATCCAGAAAGACGACGAGGTATCTACTAACAATGCAAAATGAATCTGACAG 1146
QY 1156 gacaaacatgtgtaacgataatcgtgttcaatcgtgttccacgaatccaacatcgtctaa 1215
    |||
Db 1147 GACAAACATGTGTAAAGATATCGGTTTCATCGGTTTCACAGTTCAACAAATATGCTAAA 1206
QY 1216 ctgtgtctctccaactgtgtaacatctgtcagatcgaaagcttctcgcacactgtgtgtgc 1275
    |||
Db 1207 CTGTGTCTCTCCAACTGTGTCAATCTGTCAATGTCAAGCTGTCTCTCCACACTGTGGTTC 1266
QY 1276 tcttggagatcattcccggtgtgtagtgaagctgtgggtgaagctgcgctgttaa 1326
    |||
Db 1267 TCTTGGAGATTCAATCCCGGTTGATGACGGTTGGGGTGAACGTCGCGTGTAA 1317

RESULT 4
LOCUS X022962 1338 bp DNA linear SYN 08-NOV-1995
DEFINITION Synthetic botulinum neurotoxin serotype A Hc fragment (botA) gene,
complete cds.
ACCESSION U22962
VERSION U22962.1 GI:733428
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE
1 (bases 1 to 1338)
AUTHORS Clayton, M.A., Clayton, J.M., Brown, D.R. and Middlebrook, J.L.
TITLE Protective vaccination with a recombinant fragment of Clostridium
botulinum neurotoxin serotype A expressed from a synthetic gene in
Escherichia coli
JOURNAL Infect. Immun. 63 (7), 2738-2742 (1995)
MEDLINE 95310035
REFERENCE 2 (bases 1 to 1338)
AUTHORS Brown, D.R.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-1995) Douglas R. Brown, Toxinology, U. S. Army
Medical Research Institute of Infectious Diseases, Bldg. 1425, Fort
Detrick, Frederick, MD 21702-5011, USA
FEATURES
1. 1338
source /organism="synthetic construct"
/db_xref="taxon:32630"

```

```

BASE COUNT 401 a 342 c 249 g 346 t
ORIGIN

```

```

Query Match 98.4%; Score 1311; DB 12; Length 1338;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 cgtctgctgtctacatccaatgaatatacatcaagaacatcatcatccatccatgaac 75
    |||
Db 15 CGTCTGCTGTCTACTCTTACCTGATACATCAAGAACATCATCAATFACTCTCATCTGAAC 74
QY 76 ctgcgtctagaaatccaatcaccgtgacgtctcgtcgtcgtcgtccaaatcaaacatc 135
    |||
Db 75 CTGCGCTACGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 134
QY 136 ggtcttaagttaactctgacatcgatcgaaagaatgaatgaatgaatgaatgaatgaatga 195
    |||
Db 135 GGTCTTAAGTTAACTTCTGATCCGATCGACCAAGAAATGATGATGATGATGATGATGATG 194
QY 196 tcttccaataatcgaagatcattcctgaagaatgctatcgtatatacaactatgtaagaac 255
    |||
Db 195 TCTTCCAAATCGAAGTTATCTCTGAAGATCTTCTGATATCAATCTATCTATGACAAAC 254
QY 256 tcttccaactcctctgtgacatcgatccgaataactcaactcactcctcgtgaacat 315
    |||
Db 255 TTCTCCACCTCTCTCTGATCCGATCCGAAATCAATCAATCAATCAATCAATCAATCAATCA 314
QY 316 gaataacacatcatcaactgtagtaagaacaaatctgtgtggaagaatctcgtgaactac 375
    |||
Db 315 GAATACACCATCATCACTGATGGAAGAAACAAATCTGTGTGGAAGATCTGTGAACCTAC 374
QY 376 gtgtaaatcatctgactctgcaagacactcaggaataatcaaacagcgtgtgtatcaaa 435
    |||
Db 375 GGTAAATCATCTGACTCTGACGACACTCGAGACACTCGAGAAATCAACAGCGCTGTGTCAA 434
QY 436 tactcagaatgatacaactctgtactacataatcgtcgtgtaactctgtatcaaac 495
    |||
Db 435 TACTCTGAGATGATCAACATCTGTACTATCAATCGCTGATCTGTGTACTACATCAC 494
QY 496 aacaaatcgtcgtgaataactcaaaaatcatacaaacggtcgtcgtcgtcgtcgtcgtcgt 555
    |||
Db 495 AACAAATGCTGTGAATTAATCTCAAAATGTACTACAGCGCGCTGTGATCGACCAAGAAC 554
QY 556 atctccaatctgtgttaacatcacgctcttaataacatcatgttcaaacgtggaagctgt 615
    |||
Db 555 ATCTCAATCTGTGTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 614
QY 616 cgtgacatcaacgctacatctgtgataaataactcaactcgttctgcaagaagaactgaac 675
    |||
Db 615 CGTGACATCAACGCTCACTGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 674
QY 676 gaaaaagaatcaagaacctgtacgacacacgaagtcacattctgtgtatctcgtgaagacttc 735

```

Db	675	GAAGAAAGAAATCAAGACTGTAGACACAAACCACTGCATTTCTGGTATCTCTGAAACACTTC	734
OY	736	tggggtgacctaccgcycaagtaacgaacaaaccgtaactaactgctgaaatctgtaacgtccgaac	795
Db	735	TGGGGTGCATCTACTCTCAGTACGACCAACCCGTAATCAATGCTGAAATCTGTACATCTCGAAC	794
OY	736	aaatcgtttgaagtaacgaatgtatgatactccgggttaacatgtaactgtgaaagttccggt	855
Db	735	AAATACGCTTGACGCTAACAAATGTAGGATTTCCGGGTTACATGTACCTGAAGAGTCCGGCT	854
OY	856	ggttcgttatactaccacaacatctaccctgaactctccctgtaaccgttgaacccaatlc	915
Db	855	GGTTCGTGTATGACATACCACCAATCACTGAACTCTTCCCTGTACCGTGGTACCAAAATTC	914
OY	916	atctctcaagaataatcggtctggttaacaagaacaatatcgtttcgcgaacaatgatactgtta	975
Db	915	ATTCATCAAGAAATACGCGTCTGGTAAACAAGGCAATATCGTTCGCAACATGATGTGTGA	974
OY	976	tacatcaatctgttgaattgaagaacaagaataatccgtctcggctaccacaatgctctcaagct	1035
Db	975	TACATCAATGTGTGTAGTTAAGAAACAAGAAATACCGTCTGGCTGACCAATGCTTCAAGGCT	1034
OY	1036	ggtctgaagaagaatctgtctgtcctctgtgaataatcccggaagcttgtgtaactgtctcaagta	1095
Db	1035	GGTGTAGAAAGAAATCTTGTCTCTCTCTGTGAATCCGGAGCTGGTGAATCTGTCTCAGGTA	1094
OY	1096	gtctgaatgtaaaatcccaagaacgaccgaaggtatcactaacaatagtcaaaatgatactgcag	1155
Db	1095	GTTTGAATGAATCCAAAGACGACCAAGGGTATCACTAACAAAATGCAAAATGAATGTGCGAG	1154
OY	1156	gacaaacatgtgtaacgataatcggtttcaatcggtttccacacagttcaacaataatgctaa	1215
Db	1155	GACAAACAATGGTAACGATATCGGTTTCATCGGTTTCCACCAATTTCAACAATATTCCTTAA	1214
OY	1216	ctggttgcctccaactgtgtacaaatcgtaacgtacgaacgattccctctcgacactctgggttgc	1275
Db	1215	CTGGTTGCTTCCAACTGGTACATGTGATCGAGATCGAAGCGTTCTCTCGCACTCTGGGGTTGC	1274
OY	1276	tcttggaggttcaatcccggttgaatgaacggttggggtgaaatcgccgtgttaa	1326
Db	1275	TCTTGGAGGATTCATCCCGGTTGATACCGGTTGGGGTGAACGTCGCCCTGTGAA	1325

RESULT	5					
AR000030						
LOCUS	AR000030	1402 bp	DNA	linear	PAT 04-DEC-1998	
DEFINITION	Sequence	25 from patent US 5736139.				
ACCESSION	AR000030					
VERSION	AR000030.1	GI:3962561				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 1402)					
TITLE	Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.					
JOURNAL	Treatment of Clostridium difficile induced disease					
FEATURES	Patent: US 5736139-A 25 07-APR-1998;					
Source	Location/Qualifiers					
	1..1402					
	/organism="unknown"					
BASE COUNT	420 a 360 c 260 g 362 t					
ORIGIN						

Query Match	98.48;	Score 1311;	DB 6;	Length 1402;
Best Local Similarity	100.08;	Pred. No. 0;		
Matches 1311; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 16 cgtctgctgtctaccttcactgatacatacgaacatcatcaatcctccatctgaac 75
|||||
79 CGTCTGCTGTCTACCTTCACTGATACATTCAGACACATCATCAATACCTCCATCCTGAAC 138
Db

OY	76	ctggcgtaagaaatccaatcaacacggaatcgagacggtctcgcgtatagcttccaataatcaatc	135
Db	139	CTGGCGTACGAAATCCAAATCACTACCTGATCGACTCTCGCTTCCAAAATCAACATC	198
OY	136	ggcttctaagttaactctcgatccgcatcgaaagaatccaagatccagctgtltaactcygaa	195
Db	199	GGTTCCTAAAGTTAACTCTCGATCCGATCGACGAAGAATCAATCATCTGCCTGTCAATCTGGAA	258
OY	196	tcttccaaatcgaaagttatccctggaagaatgctatcgtataaactctatgtacgaaaac	255
Db	259	TCTTCCAAAATCGAAGTTATCTCGAAGAAATGCTATGTATCAACTATATGACAAAAC	318
OY	256	tttcccaactccctcttggatcggtatccccgaatcttcaactccatctcttgaacat	315
Db	319	TTTCTCCACTCTCTTCTGGATCCGTATCCGAAATACCTTAACCTCATCTCTGTGAACAT	378
OY	316	gaatacaccatcaatcaactcgatcatgyaanaaactctggtitgtaagatctcgaactac	375
Db	379	GAATACACCATCATCACTCATGTGAAGAAACAATTTCTGTTGAAAGTATCTCTGACATAC	438
OY	376	ggtaaatcaatctggaactctgcagagacactcagaagaatcaaacagcgtgttatccaa	435
Db	439	GGTGAATCATCTTGACTCTGACGAGCACTCAAGAAATCAAAACAGGtGtGTATTCAAA	498
OY	436	tactccaagatgatacaacatctcttgataatcaatctcgttggatcttgtaaccatacc	495
Db	499	TACTCTCGATGATCAACATCTCTGCACTACATCAATACGCTGGATCTGTTACATCAAC	558
OY	436	aacaaatcgctcgtaaataactccaaaactctcaataaaggccgtctgatatcgacaagaacccg	555
Db	559	AACATCTCTCGAATTAATCTCCAAAATCTCAATCAACGGCGTGTATCCACCAAGAAACCG	618
OY	556	atctccaactctggaataacacccaacgcgtcttcaataacatcatgtltcaaatgtagcgttgt	615
Db	619	ATCTCAATCTGGGTAACTACCAACGCGTCTTAATATCATCATGTGTTAACTGGACGCTGT	678
OY	616	cgtgacactccgcgtatcatctctggaatcaatctctgttgcacaagaactgaac	675
Db	679	CGTGCACTCCACCGCTCACTCTGATCAATTAATCTTCAATCTGTGACAAAGAAACCTGAC	738
OY	676	gaaaagaagaatcaagaagccgtctacgacaacagttccaatctctgtatccctgaaagaatc	735
Db	739	GAAAAAGAAATCAAAAGACCTGTAGCAACACCATCTCAATTCGTGTATCCTGTGAAGACTTC	798
OY	736	tggggtgtaactcctctgcagtaacgaaacccgtactacatgctgaaatctgtatcgatccgaac	795
Db	799	TGGGGTGACTCACTCGCAAGTACGAACAAACCGTACTACATGCTGAATCTGTACATCCGATAC	858
OY	736	aataacgttgcgcgaacaatgtatgtaatctccgggttatactgttactgtaagttccggt	855
Db	859	AAATACGTTGACGTCACCAATATGATGATCCGGGTTACATGTACTGAAAGGTCGGGCT	918
OY	856	ggcttctgtatgtaactacaacatctactcctaactctctcccttaccggtgtgtacccaatctc	915
Db	919	GSTTCTGTTATGACTTACCACATCTACCTACCTTCCCTGTGACCGTGTGACCAATTC	978
OY	916	atcatcaagaataacgcgctcgtgtgttaacaagagacaatatcgttcgacaacaatgatacgtgtat	975
Db	979	ATCATCAAGAAATACGGCGTCTGTGTAAACAGCAAAATATGTCTTGCAAAATATATGTGTATA	1038
OY	976	tacatcaaatgtctgtactttaaagaacaaagaaataccgctcgtgtaccaaatcctctcaagct	1035
Db	1039	TACATCATCTGTTGTATGTTAAGAAACAAAGATTAACCGTCTGGCTACCAATCTCTTAGGCT	1098
OY	1036	gggtgtaagaagaatcttgtctgtcctctgaaatcccggaacggtgtgttaatctgtctcaagta	1095
Db	1099	GGGTGTAAGAAAGATCTGTCTGTCTGTGAATCCCGAGAGTGTGTATCTGTCTCAGGTA	1158
OY	1096	gttgttaagaaatcccaagaacagccagaggtatcaactaacaatgtgcaaaatgtaatctgacg	1155
Db	1159	GTTGTATGAATTCCAAGAACGACCAAGGATATCACTTAACAATGCAAAATGAATCTTGACG	1218
OY	1156	gacaacaatgtgtaacgatalcgtgttataatcgtgttccaccaggttcaacaataatcgtctaa	1215

|||||
Db 1219 GACACAAATGGTAACGATATCGTTTCATCGTTCCACAGTTCAACATATCGCTAA 1278
Qy 1216 ctgtgtcttcccaactggtacacatcgcagatgaacgcttccctcgcactctgggtgtc 1275
Db 1279 CTGGTTCCTTCAACAGTGTGATCGATCGATGGAACGTTCTCTCGCAGCTGGGGTTC 1338
Qy 1276 tcttggagatcaccggtgtgacggttggtggtgacgctccgtctgtaa 1326
Db 1339 TCTTGGAGATTCAATCCCGGTTGATGACGGTTGGGGTCAACGTCGGCTGTAA 1389

RESULT 6
ARI69141
LOCUS ARI69141 1402 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 25 from patent US 6290960.
ACCESSION ARI69141
VERSION ARI69141.1 GI:11906911
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Kink,J.A., Thalley,B.S. and Stafford,D.C.
TITLE Vaccine and antitoxin for the treatment of C. difficile disease
JOURNAL Patent: US 6290960-A 25 18-SEP-2001;
FEATURES
source 1.1402
Location/Qualifiers
BASE COUNT 420 a 360 c 260 g 362 t
ORIGIN

Query Match 98.4%; Score 1311; DB 6; Length 1402;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 cgtctgtctgtctaccccttctacatgaatgaacacatcaatctccatctgtaac 75
Db 79 CGTCTCTGTCTACTCTCAATGATATCAATCAAGATCAATCAATCAATCAATCA 138
Qy 76 ctgctgctgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 135
Db 139 CTGGCTGACGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 198
Qy 136 ggtctaaagttaactcgaatccgatcgacagaagaatcgaatccagctgttcaatc 195
Db 199 GGTCTAAAGTTAACTTCGATCGATCGACAAATCAATCAATCAATCAATCAATCA 258
Qy 196 tcttccaaaatcgaaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 255
Db 259 TCTTCCAAATCGAATGATATCGTGAAGATGCTATATCAATCAATCAATCAATCA 318
Qy 256 tcttccacccttctggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 315
Db 319 TTTCTCAACCTCTTCTTGGAATCGATCCGAATTAATCAATCAATCAATCAATCA 378
Qy 316 gaatacacatcaatcaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 375
Db 379 GAATACACATCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 438
Qy 376 ggtgaaataatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 435
Db 439 GGTGAATATATGGAATCGATCGACGACACTCAGGAATCAATCAATCAATCAATCA 498
Qy 436 tacttcagatgaatcaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 495
Db 499 TACTCTCAATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 558
Qy 496 aacaatcgtctgaataatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 555
Db 559 AACAAATGCTGTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 618

Qy 556 atctcaatctgggtgaacatccacgctctcaataacatcgttcaaacctggagctgtc 615
Db 619 ATCTCAATCTGGGTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 678
Qy 616 ctgtacacacgctacatcgtatgaatcaatcgaatcgaatcgaatcgaatcgaatcga 675
Db 679 GGTGACACTACCGCTTCAATCTGATCAATCAATCAATCAATCAATCAATCAATCAATCA 738
Qy 676 gaaagaagaatcaagaacgctgtacgaacacgctcgaatcgttgaatcgttgaagaatc 735
Db 739 GAAAAAGAAATCAAGACCTGTACGACACACGACGCAATCTGATCTGATCTGAAAGATTC 798
Qy 736 tgggtgtactactcgtacgacgaacacgctcgaatcgttgaatcgttgaatcgttga 795
Db 799 TGGGTGTACTACCTGCTGATCGACAAACCGTACTACATGATGATGATGATGATGATG 858
Qy 796 aaatacgttgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 855
Db 859 AAATACGTTGACGTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 918
Qy 856 ggtctgttctgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 915
Db 919 GGTCTGTATGACTACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 978
Qy 916 atcaatcaagaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 975
Db 979 ATCATCAAGAAATACCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1038
Qy 976 tactcaatctgtgagtgaagaagaatcgaatcgaatcgaatcgaatcgaatcgaatcga 1035
Db 1039 TACATCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1098
Qy 1036 ggtctgaagaagaatcgttctgtctgtgaatcccgagctgtgttgaatcgttctga 1095
Db 1099 GGTGTACAAAAGATCTGTCTGCTGTGGAATCCCGAGCGTGGTATCTGTCTACAGTA 1158
Qy 1096 gttgtatgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 1155
Db 1159 GTTGTATGAAATCCAAAGAACGACGATGATGATGATGATGATGATGATGATGATG 1218
Qy 1156 gacacaaatgtaacgatacgttctcgaatcgaatcgaatcgaatcgaatcgaatcga 1215
Db 1219 GACACAAATGTAACGATATCGTTTATCGTTTCCACAGTTCAACATATCGCTAA 1278
Qy 1216 ctgtgtcttcccaactggtacacatcgcagatgaacgcttccctcgcactctgggtgtc 1275
Db 1279 CTGGTTCCTTCAACAGTGTGATCGATCGATGGAACGTTCTCTCGCAGCTGGGGTTC 1338
Qy 1276 tcttggagatcaccggtgtgacggttggtggtgacgctccgtctgtaa 1326
Db 1339 TCTTGGAGATTCAATCCCGGTTGATGACGGTTGGGGTCAACGTCGGCTGTAA 1389

RESULT 7
AX036246
LOCUS AX036246 1402 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 25 from Patent EP1041149.
ACCESSION AX036246
VERSION AX036246.1 GI:11225864
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Kink,J.A., Firca,J.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and Williams,J.A.
TITLE Vaccine and antitoxin for treatment and prevention of C. Difficile disease
JOURNAL Patent: EP 1041149-A 25 04-OCT-2000;
FEATURES
source 1.1402
Location/Qualifiers
/organism="synthetic construct"

```

promoter      /note="-35 region"
              41..46
              /note="-10 region"
RBS           63..68
              /note="ribosome binding site"
              77..3967
              /note="type A neurotoxin (AA 1-1296)"
              /codon_start=1
              /transl_table=11
              /protein_id="CAA36289.1"
              /db_xref="GI:40382"
              /db_xref="SWISS-PROT:P10845"
              /translation="MQFVKKQKQYKDPVNGVDIAVLIKIPNWGOMQPVKAFKIHNTIIV
              IPERDFTNPEEGDNLNPPPEAKQVPSYSDYSTDNKDKYLVGVTLEFRISTD
              LGRMLTISVIRGIPMGSTIDTELKVIDTNCINYIOPGYSRSELTNVIIGPSADI
              IOPECKSGHEVNLTRNGYSTOYIRSPDPTPESELEEDTNPDLGAGFAPDPA
              VTLAEHLIAGHRLTGIAINPRPVKVNATYEMSGLESEBELRTGEGHAKRTDS
              LOENEFRLTYNKEFKDIASLTNKAISVGTASTLQYMKVNEFEKYLSDTSGKFSVD
              KIKFKLXKMLTEITEDNFEVFEFVNLKRYLNDKDAVKINIPKVNYTYDGFNL
              RNTNLANFNQNTIEINNNFELKLNFTGLFEFFKLLCYRGITTSKTSIDKYNKAL
              NDLCIKVNMDLFPSPEDNFTNDLNKGEITSDNIEAENISLDLQOXYLFNF
              DNEPNISLENLSDIIGOLELPMNIEPPNCKKYELEKVTMEHYLRAGEFPGKSR
              ALTNVNEALLNPSRVYTFESSDYKKVKKATKAEMAFGLGVQVLYDPTDESEVSTT
              DKIAITITIIPIGALNGLNMLYKDDEFGALIFSGAVLLEFIPBIAIPVLELTALV
              SYIAKVLIVQITDIALSKRNKEWDEVKYIVTNMLAKVNTQIDILIRKKMALENOA
              EATKAIINYOQYIEEKNININFDLSKLNESINKAMINIKFLMOCSVSYLMA
              SMIPYGVKRLPEDDASLKDALKYIYDNGCTLIGOVDRKDKVNNTLSDIDPEOLSKY
              VNOBLRSTFEYIKNIITNTSLNLRYSNHLIDSRYSKINIGSKVFPDIDKNOI
              OLPLNLESSKIEVILKNAIYVNMENESFPIRIPKYNISLNEYIITNCKMENS
              GKRVSLNTEGELIWTIADPOELKORVVFRTSOMINISDYINRAIEVTITNNRLNSKLY
              INGRILDKRPLISNLGHIHASNIMKFLDCCRPTHRYIWKYENLDEKLENEKIDLY
              DNQSGILKDFMGDYLYDKPYMLKIDYDNPKNYGVNVNNGIRGYMYLKGPSGVVTT
              NIYNSLRYGRFTLIKRYKASGNKDNINYNDRYIVNVVKKREYRLATNISOAGVEK
              ILSALEIPVGNLSOVVVKRSKNDGKITCKKMNLODNNNGNDIGFIFGFIQFNIAIKLV
              ASNMVROERSRRLGCSMERIIPVDDGGERPL"
misc_feature  4042..4087
              /note="dyad symmetry"
misc_feature  4104..4136
              /note="dyad symmetry"
misc_feature  4158..4188
              /note="dyad symmetry"
BASE COUNT  1738 a 423 c 668 g 1463 t
ORIGIN
Query Match      54.6%; Score 727.6; DB 1; Length 4292;
Best local Similarity 72.2%; Pred. No. 6.4e-177;
Matches 946; Conservative 0; Mismatches 364; Indels 0; Gaps 0;

```

```

QY 383 tcatctgactctgcagagcaactcaagaatcaacaagcgtgtgtatccaactctc 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3024 TAATCTGACTTACACAGAGTACTCAGGAATTAACAAGAGTAGTTTAAATACAGTC 3083

QY 443 agatgatcaacatctctgactacatcaatcgcctgagatctgttaccatccaacatc 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3084 AAATGATTAATATATAGATTAATATTAACAGATGATTTTGTAACTATCACATAATAA 3143

QY 503 gtcgtgaatactccaatctacatcaacagcgctgcgtgacgcgcgaagacgctccca 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3144 GATTAAATTAACCTAAATTTATATATTAATGAAGATTATAGTTCAAAACCAATTTTCAA 3203

QY 563 atctgggtaacatccacgcctctcaataacatcaatgtaacaaactggaacggtgtgaca 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3204 ATTTAGGTAAATTTATCTAGTAGTAATAAATTAATTAATGTTTAAATTAAGTGTGAGATA 3263

QY 623 ctacacgctacatctgatacaaatcaatctgctgtgacaaagaactgacgaaagaag 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3264 CACATAGATATATTTGGATTAATAATATTTTAACTTTTGTATAGCAATTAAGAAAAAG 3323

QY 683 aaatcaagaacctgtacgaacaacagtlccaactctgtaacctggaagactctggggtg 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3324 AAATCAAAAGTTTATATGATTAATCAATCAATCAATTCAGGTATTTAAAAAGACTTTGGGGTG 3383

QY 743 actaccctgcagtaacgaacaacgtaactacatgctgaatctgtacgaatccgaacaatag 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3384 ATTAATTACAAATATGATTAACCATACATATATGTTAATTAATTAATGATCCAAATTAATATG 3443

QY 803 ttgaagctcaacaatgtgaagtaatccgcggtgtacatgtgaactggaaggtccggtgtctg 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3444 TCCATGTAAATTAATGATAGATTAATGAGTATATGATCTTAAAGGGCCCTAGAGGATAGCG 3503

QY 863 tttagctaccacatctacatctgtaactctccctggtgacgtgtgtacaaatcatatca 922
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3504 TAATGACTACAAACATTTATTTAATTAATTCAGATTGTRTAGGGGCAAAAATTTATTAATA 3563

QY 923 agaataacgcgtctgtgtaacaagaacatctgcgcaacaatgtaatcgtgtatataca 982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3564 AAAAATATGCTCTCGGAAATAAAGATTAATATGTTAGAAATTAATGATCGGTATATATTA 3623

QY 983 atctgtaagtaagaacaagaatlacgcgtctgctccaaatgcttcagcgtgtgtg 1042
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3624 ATGTAGTAGTAAATAAATAAAGATTAATGATTAGCTCTAATGATCATCAGCAGCGCTAG 3683

QY 1043 aaaagatctgtctgctcgtgaaatcccgagcgttgtaactcgtctcaagtagtgtaa 1102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3684 AAAAATATGTAATGCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 3743

QY 1103 tgaatccaaagaacgacaggtgtatacaactaaatgcaaatgtaatctgcaggaaca 1162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3744 TGAAGTCAAAAATAATGATCAAGGATTAATAAATAAATGCAAAAATGATTTACAGATTAATA 3803

QY 1163 atgtgaagatcgcgtttcatcgtgttccacgaagtcaacatalcgcgtcaactgtgtg 1222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3804 ATGGGATGATTAATAGGCTTATAGGATTTATCATGATTTAATTAATTAATGATTAATGATG 3863

QY 1223 ctccaacgtgtaacatcgtcagatcgaacgttcctctcgcacatccggtgtgtcttg 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3864 CAAGTATATTGCTATTAATAGCAATAAGCAATGAATGATCTAGAGACTTTGGCTGCTCATGG 3923

QY 1283 agtcaatcccggttgatgacggttggtggtggaacgctgtaagaatc 1332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3924 AATTTATTTCTGTAGATGATGATGATGGGAGAAAGCCACTGTATTAATTC 3973

RESULT 9
CLONEUR 4835 bp DNA linear BCT 26-APR-1993
DEFINITION C.botulinum neurotoxin gene, complete cds.
ACCESSION M30196
VERSION M30196.1 GI:144864
KEYWORDS neurotoxin.
SOURCE C.botulinum (strain 62A, subtype A) DNA.

```


CDS

/db_xref="taxon:32630"
1..1389
/note="Synthetic"
/codon_start=1
/transl_table=1
/protein_id="CAC16479.1"
/db_xref="GI:11225865"
/translation="MGHHHHHHHSSGHIEGRHMSMARLSTFEYIKNIITSI
LNEVSNHLIDLSRYASKINIGSKVNDPIIDKNOIOLEFNESKIEYILKNAIYNS
MYENFSFMRIRIKYFNSISLNEEYITINENNNSSKRYLNGEITWTLODVEIK
ORVFKTSQMINISDYINRWLFVITINRNLNNSKIYINGRLIDQFISLNCNTHASNN
IMFKLDCGRDHRH IWLKFMLEFKLEKEIKDLIDQNSGILKIDGDIYLDKPK
YVMNLDPKRYVDVNNVNGIRGYMLKGRGSMVTNTNLYNSSLYRGKFTIKKVASG
NKDNIVRNDRVYINVVYKNKEIRLATNSOAGVEKILSALEIPDVGMLSOVVMKSK
NDGKITKCKMNLQDNNNGNDIGFIFGHQFNFIATLVASNMWNRQERSRLLGCSMEF
IPVDGMCERL"

ASE COUNT 420 a 360 c 260 g 362 t
RIGIN

Query Match 98.4%; Score 1311; DB 6; Length 1402;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

16 cgtctcgtctgtacctctactgaatacacaagaacatcatcaactcctcgaac 75
179 CGTGTGCTGTCTACTCTCACTGAATATCAAGAACATCATCAATCACTCCATCGTGAAAC 138

76 ctgagctacgaatccaatcaactcgtacgtctcgtcgaacgtctccaataacatc 135
139 CTGCGCTACGAATCCAAATCACCTGATCGACCTGTCTCGTACGCTTCCAAAATCAACATC 198

136 ggttcataagtaactcgaatcgaacgaacatcgaatcgaatcgaatcgaatcgaat 195
199 GGTTCATAAGTAACTTGTGATCCGATCGACAGAAATCAATCACTGCTTCATCTCGAATA 258

196 tcttcacaaatcgaagtatacctgaagaatgctatcgtatatacaactctatgaagaac 255
259 TCTTCCAAAATCGAAGTATATCTCTGAAGAATGCTATGATCACTCTATGACGAATAAC 318

256 ttctcaccctctcttgatccgataccgaaatcttaactcattcctcttgaaatc 315
319 TTCTCCACCTCTCTTGATCCGATCGATCCGAAATCTTCAATCCATCTCTTGAAACAAT 378

316 gaatacacatcatcaactcgaacgaacaaatctcgtctggaagaatcctcgaactac 375
379 GAATACACATCATCACTGATGCAAAAAATCTCTGCTGGAAGTATCTTGAAATGAC 438

376 ggtgaatcatctggaactcgtcgaagacactcagaagaatcaaacagcgtgtgtatcaaa 435
439 GGTGAATCATCTGAGACTCTGAGACACTCAGAAATCAACACGCTGTGTATTCAAA 498

436 tacttcagatgatcaaatctctgaactatacatcgtcgtgatacttgtaaccataac 495
499 TACTTCAGATGATCAAACTCTGATCAATCAATTCGCTGATCTCTGATCAATCAAC 558

496 aacaatcgtctgaacaaactcaaaatctatacatcgaacggtcgtatcgaacgaagaacg 555
559 AACATCGTCTGAATTAATCTCCAAATTAATCAATCAACGCGCTGATGACCAAGAAACGG 618

556 atctccaactcgtggaatacatcgaactcttaataacatcgttccaactggaacggtgt 615
619 ATCTCCATCTGCGTGAATCATCCACGCTTCTATTAATCATCATGTTCAAACTGAGCGTGT 678

616 cgtgacatcaaccgtatacatcgtgatacaatctcaatctgttcgacaagaactggaac 675
679 CGTGACATCAACCGGTATCATCTGATCAATCTTCAATCTTTCGACAAACACTGAAAC 738

676 gaaaagaatacaagaactctgaacacaacaggtccaatctcgtatctgaagaacttc 735
739 GAAAAGAATAATCAAGACTCTGACACACAGCTCAATCTTGATCTGGAAGAAGCTTC 798

736 tggggtgactactcgtcagtaacacaacggtactatactgtaactcgtgaacatcgaac 795

DB 799 TGGGGTACTACTGACGTACGACAAACCGTACTACATCTGTAATCTGACATCCGAAC 858
796 aaatacgtgtaacgtcaacaatagtgatccgcggtgtacatgtaactggaagtcgcgt 855
859 AATATAGGTTGAGGTCAACATGATGATGATCCGGGTACATGATCTGGAAGTCCGCGT 918

856 ggttcgtatgtaactcaacaatctcaactcgaactctcccttaccggtgtgaacaaatc 915
919 GGTTCGTATGATGATCAACATCTACTGAACTCTCCCTGTACCGGTGTGACCAATTC 978

916 atcatcaagaatacagcgtctgtgtacaaagaacatcgttcgacaacatgatactgtga 975
979 ATCATCAAGAATAATACCGCTGCTGTACAGCAAGCAATATGTTGCGAACAATGATCGTGA 1038

976 tacatcaatggtgtgtgtgaagaacaaatatacgtctgtcaccatgcttccagct 1035
1039 TACATCAATGTTGTGATGATGAAGCAAAAGATACCGTGTGCTACCAATCTTCTCAGGCT 1098

1036 ggtgtagaagaatcgtctgtcgtcgtgaatacccggaacgttgtatcgtctcagta 1095
1099 GGTGTGAGAAAGATCTTGTCTGCTGTGAAATCCCGGAGCTGTGATCTGTCTCAGGTA 1158

1096 gttgtataagaatccaagaacgacaggtgtatcactaacaatgcaaatgaaatcgtcag 1155
1159 GTTGTATGAATCCAAAGACGACAGGATATCACTAACAAATGCAAAATGAAATCTGCAG 1218

1156 gacaacaatgtaacagatcgtgttctcgtgttcaccaggttcaacaatcgatgataa 1215
1219 GACAAACAATGTAACGATTCGTTTCATCGGTTTCACCACTCAACAAATATCGGTAATA 1278

1216 cgtgtgtctcgaactcgtgataacgcagatcgaacgtctcctcgtcgaactcgtgtgc 1275
1279 CTGTTGTTTCCAACTGATCAATCTGATGATGAAAGTTCCCTGTGCACTGTGGTTGC 1338

1276 tcttggaatcattaccggtgtgatacgtgtgggtgaaacgttcgcgtgtga 1326
1339 TCTTGGAGCTTCAATCCGGTGTGATGACGGTGGGGAACGTCCTCGTGA 1389

RESULT 8
CBOTAG 4292 bp DNA linear BCT 12-SEP-1993
LOCUS Clostridium botulinum botA gene for type A neurotoxin.
DEFINITION X52066 X52088
ACCESSION X52066.1 GI:40381
VERSION
KEYWORDS botA gene; neurotoxin; secreted protein.
SOURCE Clostridium botulinum.
ORGANISM Clostridium botulinum
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.
REFERENCE 1 (bases 1 to 4292)
AUTHORS Minton, N.P.
TITLE Direct Submission
JOURNAL Submitted (08-JAN-1990) Minton N.P., PHUS Centre for Applied Microbiology & Research, Molecular Genetics Group, Division of Biotechnology, Porton Down, Salisbury SP4 0UG Wiltshire, U K
2 (bases 1 to 4292)
REFERENCE Thompson, D.E., Brehm, J.K., Oultrem, J.D., Swinfield, T.J., Shone, C.C., Atkinson, T., Welling, J. and Minton, N.P.
TITLE The complete amino acid sequence of the Clostridium botulinum type A neurotoxin, deduced by nucleotide sequence analysis of the encoding gene
JOURNAL Eur. J. Biochem. 189 (1), 73-81 (1990)
MEDLINE 90235864
FEATURES
source Location/Qualifiers
1..4292
/organism="Clostridium botulinum"
/strain="sub sp. type A, NCRC2916"
/db_xref="taxon:1491"
16..64
/note="dyad symmetry"
19..24
promoter

Db 2648 ATGAAAGTAATCATTTAATAGACTTATCTAGTATGCAATCAAAAATAATATTCGTAGTA 2707
 Oy 143 aagtaacttgcacccgctgcgaagaatcagacccgcttctcaatcttgcgaatcttca 202
 Db 2708 AAGTAATATTTGATCCATTAATAGTAATAAATCAATTTATTTAATTTAGAACTAGTA 2767
 Oy 203 aaatcgaagttatccctgaagaatgctatcgtatataactctatgacgaagaactctcca 262
 Db 2768 AAATGAGGTAAATTTTAAATAATGCTATGTATTAATAGTATGATGAATAATTTTAGTA 2827
 Oy 263 cctcccttgcgataccgtaacccgaatacttcaactcactctcttgcgaataatgataca 322
 Db 2828 CTACCTTTTGGATAGAAATTCCTAAGTATTTTAAACAGTATTAAGCTTAATATGAAATATA 2887
 Oy 323 ccatacgaactgcgatagaagaagaatcttggttggaaagatctctgacacgagtgaa 382
 Db 2888 CAATTAATTAATTTGATGGAATAATTAATTCAGATGGAAGATACCTTAATTTATGCTGANA 2947
 Oy 383 tcaatctgactctgcaggaacacacgaagaatcaacaacgctgctgattcaataactctc 442
 Db 2948 TAACTGCACTTTACAGAGTAAGTCTCGAATAAATAAAGAGTACTTTTAAATACAGTC 3007
 Oy 443 agatgatacaacatctctgatacacaatcgcgtgagatctctgattacacacacaatc 502
 Db 3008 AAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3067
 Oy 503 gtcgaataactccaactcacaacacgagccgctgctgacacgaagaacacgactcca 562
 Db 3068 GATTAAATTAATCTTAATAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 3127
 Oy 563 atctgggtaacatccacgctctcaataacatcatcttcaactgcgagctgctgaca 622
 Db 3128 ATTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3187
 Oy 623 ctccacgctacatctgatacacaatcactcaatctgctgacacgaagaactgaaacaaag 682
 Db 3188 CACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3247
 Oy 683 aaatcaagaactctgacacacacgctcactcactgctgatacctgtaagaactcttggtg 742
 Db 3248 AAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3307
 Oy 743 actacccctgacgatacacaacacgctcactcactgctgatacctgtaagaactcacaac 802
 Db 3308 ATTATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3367
 Oy 803 ttgacgctcaacaatgtaagtaacgctgctgacatgtaacgtaagaactgctgctgctg 862
 Db 3368 TCGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3427
 Oy 863 ttaagcctacacacacacacgctcactcactgctgatacctgtaagaactcacaac 922
 Db 3428 TAAATGACTACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3487
 Oy 923 agaaatagcgtctgtaacacacacacacacacacacacacacacacacacacacacacac 982
 Db 3488 AAAATATATGCTTCGGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3547
 Oy 983 atgttgaagtttaagaacacacacacacacacacacacacacacacacacacacacacac 1042
 Db 3548 ATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3607
 Oy 1043 aaaagatctgctgctgacacacacacacacacacacacacacacacacacacacacacac 1102
 Db 3608 AAAAATATCTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3667
 Oy 1103 tgaatcacaagaacacacacacacacacacacacacacacacacacacacacacacacac 1162
 Db 3668 TGAAGTCAAAAATGATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3727
 Oy 1163 atgttgaacacacacacacacacacacacacacacacacacacacacacacacacacac 1222

Db 3728 ATGGGAATGATATAGGCTTTATAGGATTTTCACTAGTTAATAATATAGCTAAACATAGAG 3787
 Oy 1223 ctcccaactggtacacatctgcagatgaacacacacacacacacacacacacacacacac 1282
 Db 3788 CAAGTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3847
 Oy 1283 agttaccccggttgcagcgttgcggtgacacacacacacacacacacacacacacacac 1326
 Db 3848 AATTTATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3891

RESULT 12
 AX036248
 LOCUS AX036248 3891 bp DNA linear PAT 16-NOV-2000
 DEFINITION Sequence 27 from Patent Epi041149.
 ACCESSION AX036248
 VERSION AX036248.1 GI:11225866
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 artificial sequence.
 1 (bases 1 to 3891)
 REFERENCE
 1 Kink,J.A., Filica,J.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and
 AUTHORS Williams,J.A.
 TITLE Vaccine and antitoxin for treatment and prevention of c. Difficile
 disease
 JOURNAL Patent: EP 1041149-A 27 04-Oct-2000;
 FEATURES
 source
 location/Qualifiers
 1..3891
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..3891
 /note="synthetic"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC16480.1"
 /db_xref="GI:11225867"
 /translation="MOFVNKOFNKDPVNGVDIAVIRKIPNYGOMOPVAKFIHINKIWI
 IPEDPTFNPREGDLNPPPEAKOPVSYSTDYSTLSTDEKKNYILAKGYKLEPRITST
 IGRMLTISIVGIPWGGSTIDTELKVIDTNCINVTODGSTRSEELVLITIGPADT
 IOPDKSPGHVNLITRNGYSTGVIRSPDPTGFESELEVDNPLIGAGKFAFDPA
 VTLAHELHACHRLVGLAIDNPNRVKVTNAYVEGSGLEVEFEELRTGCHADAFIDS
 LOENEFRLYYNKKFKDIASTLNKASISYGTASLOKYNVREKYLISDTSKESVD
 KIKFDKIKMLTEIYTEEDNPFVKFRKRTYINDKRFAFKINPKVNYIYGFNL
 RNTNLAENFONTEINMNTFKLKNFGLIEEYKLLVREGIITISKYSLDKYKAL
 NDLCIKVNNMOLFSPSEDNFTNDLNKGEELITSDNIEAEBENISLDLQIYLYTFNF
 DNEBENISIEMLSDIIGLELMPNIEEPNGKVEIDKRYMFHILRQDEHGSRI
 ALTNVSENLALNPSRVYFFSSDYVKKNKATEAMFLGYEOLVYDPTDETSVST
 DKIDITIIPIYIGPALNIGMLKDEVDGVALIFSGAVILLEFIEAIPVIGFALV
 SYANKVLTVOITNALSKRNEKMDVEYKLVITNMLAVNTOIDLIRKKMEALENOA
 EATKALINVOYNOTTEBEKNININRIDLSKINISIKKAMININIKPIINOSYSILM
 SWITPGVRLDEFDASLDALCKIYDNRGLIGVDLKLKRVNNTLSTDIPOLSKI
 VDNORLSTFEYIEIKNIINTISILNIRESNHLIDISRAKINIGSKVNFDPIDKNOI
 QUNLESSEKIEVILKNAIYVNSMENEFTSMIRIPKXFNSISLNEETIINCNENS
 GKVSLNAGEIISLLODOIEKORVFEYSOMINISDYINMIEVYITNNBLNSKIY
 INGRILDOKPIISNIGNIHASNNIMEKLKGDGTHRYIWKYENLEDEKLKREKMDY
 DNOASNGILKDFMDGYLOYDKPYMLNIDYDNKYVDVNNVGIRGMVYIKGPRJMTT
 NITYLNSISYRGTKEFIKKYIYASGNKDNIYRNNDRYINIVYVKKERELATNNSQAVEK
 ILSAELTIDVGNLSOVVVKSRNDGKITNKKMNIQDNNGDIGIFGHQHNINIAKLIV
 ASNMVNRDIESSRILGCSWEFIPVDDGWBRLP"

BASE COUNT 1580 a 392 c 629 g 1290 t
 ORIGIN

Query Match 54.5%; Score 726.4; DB 6; Length 3891;
 Best Local Similarity 72.3%; Pred. No. 1.3e-176;
 Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;
 Oy 23 tgtctaccttcaataataacacacacacacacacacacacacacacacacacacacacacac 82
 Db 2588 TATCTAATTTACTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAGAT 2647

[illegible]

Db	3728	ATGGAAATGATATGACCTTTATAGATTTTCATTCAGTTTAATAATATCTACTGAACATGAG	3787
Oy	1223	ctccaactggacaaatgcacagacgaacgttccctcgcgaactctgggtctctctgg	1282
Db	3788	CAAGTAATTTGGTATTAATAGACAATAAGAAATGATCATTAGAGACCTTTGGCTTGCTCATGGG	3847
Oy	1283	agttcatcccggttgatagcaggtctggggctgaacgctcgctgtaa	1326
Db	3848	AATTTATTCCTGTAGATGATGATGGGGAAGAAAGGCACCTGTAA	3891

RESULT	13
CANOTOXA	
LOCUS	CBNT0XA 4067 bp DNA linear BCT 24-JUN-1994
DEFINITION	C.botulinum gene for infant neurotoxin type A.
ACCESSION	X73423
VERSION	X73423.1 GI:507070
KEYWORDS	botulinum neurotoxin; botulinum neurotoxin type A.
SOURCE	Clostridium botulinum.
ORGANISM	Clostridium botulinum.
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.
AUTHORS	1 (bases 1 to 4067)
TITLE	Williams A., East A.K., Lawson P.A. and Collins M.D.
JOURNAL	Sequence of the gene coding for the neurotoxin of Clostridium
MEDLINE	botulinum type A associated with infant botulism: comparison with
REFERENCE	other clostridial neurotoxins
AUTHORS	Res. Microbiol. 144 (7), 547-556 (1993)
TITLE	94143603
JOURNAL	2 (bases 1 to 4067)
MEDLINE	Williams MCG A.
REFERENCE	Submitted (17-JUN-1993) A. Williams MCG, AFRC Institute of Food
AUTHORS	Research, Dept of Microbiology, Reading Laboratory, Earley Gate,
TITLE	Whiteknights Road, Reading, RG6 2EF, UK
JOURNAL	Location/Qualifiers
FEATURES	
SOURCE	1. 4067
	/organism="Clostridium botulinum"
	/strain="Kyoto-F"
	/db_xref="taxon:1491"
	83. 88
	97. 3987
	/codon_start=1
	/transl_table=11
	/product="botulinum neurotoxin type A"
	/protein_id="CA51824.1"
	/db_xref="GI:507071"
	/db_xref="SPRMBL:045894"
	/translation="MPYNNQFNTKQDVNGVDIAYIKIPNAGQMPPKAKFKHKIKWV
	IPEKDTNPDEGLNPPPEKQVPVSYOSTYISTONEKNTLKGVKLPERYSTD
	LGRLRLSIVRGIPEFWGSDITDELAKIDNCLINWIDPDSYSEELNLVIGSADID
	IOEGCSFGHSDVNLTRNGSGSTQIFRSPDFGPEFSELEVDNPILGAGKFTDPA
	VTLAHELIAHREHYGJAINDPNREKNTANAYMSGLSEFEELPRFEGHDAFIDS
	LOEKEPFLYYHNRKFDVASTLNKAKSLIGTGTASQWKKNNPKKEYLLSEPTSGESVD
	KLEKDKLYKMLTELYTEBDNVNFPKVIYNKTYLNFDAVRINIVPEBNTIKGFNL
	KGALSTNFGQNTLEINSRNFTRLKNFTGLFEFYKLLCVAGIIPFKYSLDEGYNKL
	NDLICILVNMNDLEFSPSEDEFTNDLVEEITDADTNEAEENSLDIOOYLYTPPE
	DNEPENSISENLSDIIGOLEPMPNIEFRPFGKKYELDKTDFMFLYLADEEHSRIR
	ILRNSAEALILKPNVATYFPFSKYVKKINKAAVEFMLMAEELVYDFPQETNEVTVM
	DKIADITIIYPIYIGPALNIGMLSGFEVALIIFTGVAMLEPIPEVALYEGFALIV
	SYLANKLVVQTIINNALSKRNEKMDVEYKTYTMMIAKLVNTQDLIEKKMALENOA
	EAKRAIINYOYNTTEEEKNNINFINIDLSLTKNESINSAMININRKLDCQSVYLN
	SMPIPAVKRLKDPASVADVLAKIYDNRGTLVQVDRKDEVAENTLSADIPQLSKYV
	VDMKILSTFEYIKNIVNTSILSIYKKDILLDSRYGAKINIGDGVYSDIKNDID
	KLINLESSTIEVILKNAIVYNSMTENSTSPWIKIPKYSKININNTYIIINCENNS
	GKWSLVSGEILWTLQDNKQIQRVFKIGQMVNISDIYINRMIEVYTTINRLTKSLY
	INGLLDQKPIISNLGNHASNKIMFKLDGCRDPRKYIMVYFNLDEKLEKEIKDILY
	DYSQNSGILKDFPMGNLQYDKRPYMLNLPDKKVDVNNYIGIRGYMLKRGSGVYTT
	NIYSNLSLEGTEKFEIIRKYVSGNEDNIVRNNDRYINVVAKREYRLATNAGSGVGEK
	ILSLAETIPDGNLSQVYVMSKDDOGIRNCKMNLQDNNGNDIGFIGPHLYDNTAKLV
	ASNMYNRQVGRKASRSTFGCCSMELPIPVDDGKESSL"
	1364. 1386

misc.feature

[illegible]

QY	923	agaatagcgtctggtgaacagaacaaatcgtctgcgaacaatgagtctgtataacatca	982
Db	3564	AAAAATATGCTTCTGGAATGAAGATTAATTTGTTAGAAATATGATGCTGTATATATTA	3643
QY	983	atgtgtatgtaagaacaagaataacccgtctgcgtaccacatgctctcaagctgtgtag	1042
Db	3644	ATGTAGTAGTTAAAAATAAAGATATATAGCTTAGTACTAATATGCATACAGCAGCGCTAG	3703
QY	1043	aaaagatctgtctgtctctgtgaaatcccgagacgttgttaactgtctcgaagtagtgttaa	1102
Db	3704	AAAAAATCTAGTAGTGCATATTAGAAATACCTGATGTGTGGGAAACTAGTCAAGATGATGAA	3763
QY	1103	tgaatccaagaacagccaggtctcactaaacaatgtcaaaatgtaactgtgaagcaaca	1162
Db	3764	TGAAGTCAAAAGATGATCATACAGGAATMAAATAATTAATGCAAAATGAAATTTACAGTAATA	3823
QY	1163	atgtgaacgatacgtgttccacgtgtccaccagttcaacaatacgtcctaactggttg	1222
Db	3824	ATGGGAATGATATAGAGCTTTATAGGATTTTCATTTGTATGATATATATAGTAAACATGATAG	3883
QY	1223	cttccaaactgtgtacaatcgttaagatcgaaagtcctctcgcgaactctgtgtgtctctg	1282
Db	3884	CAAGTAATTTGGTATTAATATGACAAAGTGGGAAAAAGCATAGTAGGACTTTCCGTTGTTTCATGGG	3943
QY	1283	agttaccccggtgtgaacggtgtgtgtgtgaacgtcgcgtgttaagaatc	1332
Db	3944	AGTTATATCTGTATGATGATGATGAGGAGAAAGTTCACTGTATTAATATC	3993
RESULT	14		
LOCUS	AF251281		
DEFINITION	Synthetic construct recombinant botulinum toxin F HC fragment gene,		
ACCESSION	AF251281		
VERSION	AF251281.1		
KEYWORDS	synthetic construct,		
SOURCE	synthetic construct		
ORGANISM	artificial sequence.		
REFERENCE	1 (bases 1 to 1299)		
AUTHORS	Holley,J.L., Elmore,M., Mauchline,M., Minton,N. and Titball,R.W.		
TITLE	Cloning, expression and evaluation of a recombinant sub-unit vaccine against Clostridium botulinum type F toxin		
JOURNAL	Vaccine 19 (2-3), 288-297 (2000)		
MEDLINE	20389730		
PUBMED	10930684		
REFERENCE	2 (bases 1 to 1299)		
AUTHORS	Holley,J.L., Elmore,M., Mauchline,M., Minton,N. and Titball,R.W.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-APR-2000) Biomedical Sciences, CBD Porton Down, Salisbury, Wiltts SP4 0UQ, UK		
FEATURES	Location/Qualifiers		
source	1..1299		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
CDS	1..1299		
	/codon_start=1		
	/transl_table=1		
	/product="recombinant botulinum toxin F HC fragment"		
	/protein_id="AA001403.1"		
	/db_xref="gi:9885329"		
	/translation="MSYTNDKILILYFNKILYKKIKDNSILDMRENNKFKIDISGYSNS		
	ISLNGDYIYSTRNNGIYSSKSEVNIQANDIITNGRONSISFVWATPKYFPNKK		
	VNLNNEVITIDICIRNNNSGKTSILNKKIWTLDITDIGNNOKLFTNTOMTISIDYIN		
	KMIWVITNNKILGNSRIYINGNLIDKSIWGLIHVSDNLIFFIVCNDRIYVIGIR		
	FKVDELTGKTEITFLYSDPDPSILIDPFMGNYLLYNNKRYLLMLRTDKSITONSNE		
	LNINQAGVYOKPNIFSNTRLYTGVETVYIIRKNGSTDISDNEFVRKNDLAYINAVDR		
	VEVLIADGISTAREKIKILIRTSNNSNGIILVMSIGDNCIMNPNONNNGIGILL		
	GPHSNNLVASSWYNNIRKNTSSNGCWSFISKHGMDEN"		
BASE COUNT	397 a 342 c 207 g 353 t		
ORIGIN			

OY	1038	tgtagaanaagacatcgtctgcgtctgaaatcccgagcttgtaa---tcgtctcagat	1094
Db	1041	ACCGGAAAAAATCATCAACATCATCCGTTACTCTAACTTAAACAACCTCTCTGGGTACAGAT	1100
OY	1095	agtgtaatgaataatccaagaacgaccaggtatctcaactacaatgcaaaatgaatctgca	1154
Db	1101	CATCGTTATGGACTC-----GATCGGTAAACAATCAGATTAACACTTCCA	1145
OY	1155	gagacacaaatgtaacgatatcggtttcagtttcacacacagttcacaacatctgctaa	1214
Db	1146	GAAACAACAACGGTGGTAACATCGGTCTGCTGGGGTTTCCACTTAACAAAC-----	1194
OY	1215	actggttgcctccaactcggtacacatcgttcagaaagcttcctctgcacactcggatg	1274
Db	1195	-CTGGTTGCTTTCTTTGGTACTTAACAACAACATCGTAAAAAACACTTTCTTAACGGTTG	1253
OY	1275	ctctctggagatcaccccggtgtgatcgaagcttgggggttaa	1314
Db	1254	CTTCTGCTTTTCATCTCTAAAGAACAACGGTGGCGAAGAA	1293

Query Match	31.7%;	Score 421.6;	DB 6;	Length 1313;
Best Local Similarity	62.2%;	Pred. No. 5,8e-98;		
Matches	809;	Conservative	0;	Mismatches 433; Indels 57; Gaps 7;

QY	30	cttcacgtgaatacacaagaacatcaataacatccatccacatccgtgaacctgacgaatc	89
Db	44	CTTCACAACAACGTGTACAAAAAATAAAGACACATCTATCCGGACATCGCTTACGAAA	103
QY	90	caataactgatcgacccgtctcgctacgcttccaaatacaacatcggtttcgaagttaa	149
Db	104	CAACAATTCATCGACATCTCTGGGTATGGTTCCTTAACATCTCATCAACGGGACGCTTA	163
QY	150	cttgcataccgatcgacaagaatccagatccagctgttcaatctcggatcttccaaatcga	209
Db	164	CATCTACTCTATTAACCCGACACAGCTGGGTATCTTACTCTTTAAACCGTCTGAAGTAA	223
QY	210	agttatctcgaagaatgctatctcgtatatacaactctatglaagaaactctccactcct	269
Db	224	CATCGCTCAGAACACGACATCATCTTACAACGGTGGTATCGAGACTCTCTATCTCTTT	283
QY	270	cttgcataccgatcccggaataacttcaactcctctcttgaacatgatatcacatcat	329
Db	284	CTGGGTGTGTATCCCGGAATACTTCAACAAAAGTTAACTGAAACAACGATACATCAT	343
QY	330	caactcgat---ggaanaacatctcgtcttggaagaatctatcttgaacttcggttgaatcat	386
Db	344	CGATCGATCCGTAAACAACAACACTCGGTGTGAAAATCTCTCTGAACTCTCAACAACAATAT	403

OY	387	ctggacccttcgaagaaactcgaagaatcaaaacggctgctgatacctaataatctccgaat	446
Db	404	cttgacactctcgaggacactcctctgtaacaaacgaagaaactcggcttttcaactctacactcaaat	463
OY	447	gatacaatctctcgaatactacaatcgcgtgatacttcgtttacacatcaacaacatcgcct	506
Db	464	gactctatctctgactgacactttaaataatgaaatcttcggtactatcacttaacaaacgcct	523
OY	507	gaataactcaaaaatctacaatcaaaagcgccgtctgatacgacgaaaacgcgatcccaatc	566
Db	524	gggttaactctcgatctactacatcaacggttaacctgatcagatgaaaaatattctcttaacct	583
OY	567	gggttaaacatccacgctcttcaataacataatgttcaaacgagcgctgttcgtgacacca	626
Db	584	gggtgaatcacacggtttctgacaaacatccgtttcaaaaactggtgtgtccaaacgacac---	640
OY	627	ccggtacatcttgatcaaatactccaatcgttctgacaaagaactcgaacgaaaaaaat	686
Db	641	gcggttactggttgatccggttacttccaagttttcgacactgaaactcgggttaaaactgaat	700
OY	687	caaaagccctgacgaacaaacagtcacaaatctcgtgatacctgaaagaacttcctggtgta	746
Db	701	cgaaacctctgactgactgacaaacacggagaccgctgattccgaaagaactctcgggtaacta	760
OY	747	ccctgcagtaagaaaacccgactacaatctgtatgaatctgtacgatccgacaataatcgttga	806
Db	761	ctctctcttacaacaaacggttacttacttccgtgaacactcgtcccgatgacaaatctatcac	820
OY	807	cgtcaaaactgatactccggttatactgatactgacccaagaagctccgctggttctgatat	866
Db	821	tcgaacactct-----aaacttcctgaacatcaaacacgacagcgtgggtttatca	868
OY	867	gactaccacatctacactgaactccttcctctgataccggtgtacaaatcattcaatcaaga	926
Db	869	gaaaactcaaatattcttcttcaaacactcgtctgtaacactggtgtgaagttaattcatctcgtaa	928
OY	927	atacgcgtc-----ggtacaagaagaataatcgtctgcacaacatgatactgata	977
Db	929	aaatggttctactgacatctcttcaacactgacactcgtgactgaataaaacgacctggctta	988
OY	978	catcaatgtctgataagacaacaaagaataacgctcgtgcatacaaatgtcttcagagctcg	1037
Db	989	catcaacggttctgtagacggtgacgttgaaatnccgctgtgacgtgacatctctatnccgttaa	1048
OY	1038	tgtgaagaagatcttgcctcgtcctctggaatcccgagcgttgtaa---ctgcctcaggt	1094
Db	1049	accggaaaaaatatcaaaactgactccgtacttcaactcttaacaactctcgggctacat	1108
OY	1095	agttgtataaataatcaagaagcagcaggatctacataaataatgaatgaatctgca	1154
Db	1109	catcgttatatgactc-----gactcggtaacaaactcactatgaattctca	1153
OY	1155	ggaacaacatggtlaacagatacgttctacatcgtttccacacagttccaacaataatcgttaa	1214
Db	1203	-ctggtgtgtcttcttctgtgactacaacaaacatccgttaaaacactcttcttcaacggttg	1261
OY	1275	ctcttggaagtatacccggtgtgatacaggtctgggtgtgaa	1314
Db	1262	cttctgtggtcttcttcaataaacacggtgtggcagga	1301

THIS PAGE BLANK (USPTO)

us-09-611-419a-1.rng

3332	22	AAA54588	Sequence encoding
5546	19	AAV30575	Clostridium botuli
835	21	AAC64582	BoNT/A neurotoxin

Sequence encoding
Clostridium botuli
BoNT/A neurotoxin
DNA encoding nativ
DNA encoding nativ
C. botulinum type

Page 1
Sach

Page 1

Page 1

Page 1

Page 1

Page 1

ance

2

0


```
XX OS Synthetic.
XX FH Key
XX FT CDS
XX FT 1.1317
XX FT /tag= a
XX PN W09612802-A1.
XX PD 02-MAY-1996.
XX PF 23-OCT-1995; 95WO-US13737.
XX PR 07-JUN-1995; 95US-0480604.
XX PR 24-OCT-1994; 94US-0329154.
XX PR 16-MAR-1995; 95US-0405496.
XX PR 14-APR-1995; 95US-0422711.
XX PA (OPHI-) OPHIDIAN PHARM INC.
XX PI Firca JR, Kink JA, Padhye NV, Stafford DC, Thalley BS;
XX PI Williams JA;
XX DR WPI. 1996-230603/23.
XX DR P-PSDB; AAR95008.
XX PT Fusion proteins comprising non-toxin protein and part of toxin
XX PT useful to form anti-toxins against Clostridium botulinum type A, and
XX PT C. difficile type toxins, and to treat C. difficile intoxication,
XX PT partic. diarrhoea
XX PS Example 22; Page 336-38; 434pp; English.
XX CC A synthetic gene (AAR95008) codes for the heavy chain C fragment
XX CC (AAR95008) of Clostridium botulinum type A neurotoxin (see also
XX CC AAR95010). Codon usage allowing efficient gene expression in Escherichia
XX CC coli was utilised. The gene in vector pAlterBot was used to make
XX CC expression constructs in which fragments of C. difficile toxin A
XX CC repeat domains were expressed as genetic fusions with the C.
XX CC botulin C fragment and expressed in E. coli.
XX SQ Sequence 1330 BP, 400 A; 339 C; 246 G; 345 T; 0 other;

Query Match 98.4%; Score 1311; DB 17; Length 1330;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 cgtctgctgtctactcttccatgaatacatcaagaacatcatcaatcaatcctcctgaac 75
DB 7 cgtctgctgtctactcttccatgaatacatcaagaacatcatcaatcctcctgaac 66
OY 76 ctgctgctgaatccaatctactgtgacccgtctcgttaagcttccaaatcaatc 135
DB 67 ctgctgctgaatccaatctactgtgacccgtctcgttaagcttccaaatcaatc 126
OY 136 gcttcaagaatcctgcatcgatcgacgaagaacatcgatcagctgttcaatcggaa 195
DB 127 gcttcaagaatcctgcatcgatcgacgaagaacatcgatcagctgttcaatcggaa 186
OY 196 tcttccaaaatcgaagtatctcctgaagaatgtctatcgtatataactctatgtacgaanaac 255
DB 187 tcttccaaaatcgaagtatctcctgaagaatgtctatcgtatataactctatgtacgaanaac 246
OY 256 ttcttcacctcttctgtgctcgatccggaataactccaatcctcctgaacat 315
DB 247 ttcttcacctcttctgtgctcgatccggaataactccaatcctcctgaacat 306
OY 316 gaatacaccatcatcaacgcatggaanaacaaattcgtgtggaagtatcctgaactac 375
DB 307 gaatacaccatcatcaacgcatggaanaacaaattcgtgtggaagtatcctgaactac 366
OY 376 ggtgaatacatctgactctgacgacactcaggaaatcaacaacgctgttattacaa 435
```

```
DB 367 ggtgaatacatctgactctgacgacactcaggaaatcaacaacgctgttattacaa 426
OY 436 tactctcagatgatcaaacatctctgactatcaaatccgttgtatctgtttccatcac 495
DB 427 tactctcagatgatcaaacatctctgactatcaaatccgttgtatctgtttccatcac 486
OY 496 aacaatcgtctgaataactccaanaatctacatcaacgctgtgtatcgaaccagaacg 555
DB 487 aacaatcgtctgaataactccaanaatctacatcaacgctgtgtatcgaaccagaacg 546
OY 556 atctccaatctggtgaataatccaacgcttctaataacatcatgttcaactggaagctgt 615
DB 547 atctccaatctggtgaataatccaacgcttctaataacatcatgttcaactggaagctgt 606
OY 616 cgtgacactcacccgctacatctgtgatacaatcttaactctgttcgaanaagaactgaac 675
DB 607 cgtgacactcacccgctacatctgtgatacaatcttaactctgttcgaanaagaactgaac 666
OY 676 gaaaagaatacaagaactgtacgaacaacccagttccaattcgttactcgaagaacttc 735
DB 667 gaaaagaatacaagaactgtacgaacaacccagttccaattcgttactcgaagaacttc 726
OY 736 tgggtgacttactcgtcagtaagcaaacccgtactacatgtctgaatctgtacgaatccgaac 795
DB 727 tgggtgacttactcgtcagtaagcaaacccgtactacatgtctgaatctgtacgaatccgaac 786
OY 796 aataagctgagtaagaataatgtatgtatccgggttgaatgtactgtaagaagctcggt 855
DB 787 aataagctgagtaagaataatgtatgtatccgggttgaatgtactgtaagaagctcggt 846
OY 856 ggtctgttactgactccaacatctccggaactctccctgtacccgtgtgatacaattc 915
DB 847 ggtctgttactgactccaacatctccggaactctccctgtacccgtgtgatacaattc 906
OY 916 atcatcaagaanaatcagcgtctgtgttaacaaaggacaatatcgttcgcaacaatgactgtga 975
DB 907 atcatcaagaanaatcagcgtctgtgttaacaaaggacaatatcgttcgcaacaatgactgtga 966
OY 976 tacatcaatgtgtgtaagtaagaacaagaataaccgctgtgtccacaaatgtcttcagct 1035
DB 967 tacatcaatgtgtgtaagtaagaacaagaataaccgctgtgtccacaaatgtcttcagct 1026
OY 1036 ggtgtgaagaagaactctgtctgtctgtgaatcccggaactgtgtatcatcgtctcagga 1095
DB 1027 ggtgtgtgaagaagaactctgtctgtctgtgaatcccggaactgtgtatcatcgtctcagga 1086
OY 1096 gttgttaatgaatccaagaacgacagggatcatcaactaacaatgtcaaatgtatcgtcag 1155
DB 1087 gttgttaatgaatccaagaacgacagggatcatcaactaacaatgtcaaatgtatcgtcag 1146
OY 1156 gacaacaatgtaagaatctcgtttcattcattcattcattcattcattcattcattcattcattc 1215
DB 1147 gacaacaatgtaagaatctcgtttcattcattcattcattcattcattcattcattcattcattc 1206
OY 1216 ctggtgtcttccacactgcatcgttcagatcgtaacgtttcctctcgtcactcgtgggtgc 1275
DB 1207 ctggtgtcttccacactgcatcgttcagatcgtaacgtttcctctcgtcactcgtgggtgc 1266
OY 1276 tcttgggaatccaatcccggttgaatgagcttgggtgtgaacgttcgctgttaa 1326
DB 1267 tcttgggaatccaatcccggttgaatgagcttgggtgtgaacgttcgctgttaa 1317

RESULT 3
AAV30571
ID AAV30571 standard; DNA; 1330 BP.
XX
XX AAV30571;
AC AAV30571;
DT 07-DEC-1998 (first entry)
XX
XX Clostridium botulinum toxin A fragment C gene in pAlterBot.
```

```

XX AntiToxin; vaccine: neurotoxin; toxin A; intoxication; immunogen;
KW botulinum; ds.
XX Clostridium botulinum serotype A.
OS
FH Key Location/Qualifiers
FT CDS 1..1317
FT misc_difference 1..6
FT /tag= a
FT /tag= b
FT /note= "PALTER vector-derived nucleotides
          (encode Met-Ala)"
XX WO9808540-A1.
XX 05-MAR-1998.
XX
XX 28-AUG-1997; 97WO-US15394.
XX
XX 28-AUG-1996; 96US-0704159.
XX
XX (OPHT-) OPHIDIAN PHARM INC.
XX
XX Thalley BS, Williams JA.
XX
XX WPI; 1998-230234/20.
XX
XX P-PSDB; AAM68389.
XX
XX Host cell containing recombinant expression vector encoding
XX Clostridium botulinum type B or E toxin - useful to treat humans
XX and other animals at risk of intoxication with clostridial toxin
XX
XX Example 22; Page 262-263; 428pp; English.
XX
XX This is the DNA sequence of the Clostridium botulinum serotype A
XX toxin C-fragment gene contained in plasmid palterBot. Recombinant
XX C-fragment proteins have been produced in Escherichia coli as
XX fusion proteins with either maltose binding protein or
XX Clostridium difficile type A toxin (see AAM68387). The invention
XX relates to recombinant proteins derived from C. botulinum toxins.
XX Methods are provided which allow for the isolation of soluble
XX recombinant proteins free of significant endotoxin contamination.
XX Preferred hosts for production of recombinant proteins are E. coli,
XX insect cells and yeast cells. The recombinant toxin proteins are
XX used as immunogens for the production of vaccines and antitoxins
XX that are useful in the treatment of humans and animals at risk of
XX intoxication with clostridial toxin.
XX
XX Sequence 1330 BP; 400 A; 339 C; 246 G; 345 T; 0 other;
SQ

```

```

Query Match      98.4%; Score 1311; DB 19; Length 1330;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 cgtcgtctgtctaacctcactgaatacatcaagaacatcatcaatactcctcgaac 75
DB 7 cgtcgtctgtctaacctcactgaatacatcaagaacatcatcaatactcctcgaac 66
QY 76 ctgctgctgaatcaatcaatcctgctgacctgtctcgtcaagcttccaaaatcaatc 135
DB 67 ctgctgctgaatcaatcaatcctgctgacctgtctcgtcaagcttccaaaatcaatc 126
QY 136 ggtctcaaatgaatcgtatcgatcgacgaagaatcagatccagctcgttcaatcgtgaa 195
DB 127 ggtctcaaatgaatcgtatcgatcgacgaagaatcagatccagctcgttcaatcgtgaa 186
QY 196 tcttccaaaatcgaaagtatctcgtgaagaatgctatcgatatacaactctatgtaacgaac 255
DB 187 tcttccaaaatcgaaagtatctcgtgaagaatgctatcgatatacaactctatgtaacgaac 246
QY 256 ttctccaactcctctgtgactcgatcccgaaataacttcaactccatctctcgaacat 315

```

```

DB 247 ttctccaactcctctgtgactcgatcccgaaataacttcaactccatctctcgaacat 306
QY 316 gaatacacatcaatcaatcgtatcgatgaagaagaatcctcgttgaagaatctcgtgaactac 375
DB 307 gaatacacatcaatcaatcgtatcgatgaagaagaatcctcgttgaagaatctcgtgaactac 366
QY 376 ggtgaagaatcgtatcgatcgtcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 435
DB 367 ggtgaagaatcgtatcgatcgtcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 426
QY 436 tacttccagatgatacaatcctcgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 495
DB 427 tacttccagatgatacaatcctcgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 486
QY 496 aacaatcgtctgaataatcaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 555
DB 487 aacaatcgtctgaataatcaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 546
QY 556 atctccaatcgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 615
DB 547 atctccaatcgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 606
QY 616 cgtgaacatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 675
DB 607 cgtgaacatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 666
QY 676 gaagaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 735
DB 667 gaagaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 726
QY 736 tgggtgactacatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 795
DB 727 tgggtgactacatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 786
QY 796 aatacgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 855
DB 787 aatacgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 846
QY 856 ggtctgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 915
DB 847 ggtctgttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 906
QY 916 atcatcaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 975
DB 907 atcatcaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 966
QY 976 tacatcaatgcttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1035
DB 967 tacatcaatgcttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1026
QY 1036 ggttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1095
DB 1027 ggttgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1086
QY 1096 gttgttaatgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1155
DB 1087 gttgttaatgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1146
QY 1156 gacaagaatgtaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1215
DB 1147 gacaagaatgtaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1206
QY 1216 ctggttgcctcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1275
DB 1207 ctggttgcctcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1266
QY 1276 tcttggagatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1336
DB 1267 tcttggagatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatcgaagaatc 1317

```

RESULT 4

Db 1095 gtgtatgaatccaagaccagcaggtatcactacaatcgaatgaatctgca 1154
 QY 1156 gacacaaatgaacgatcatcggttcatcggtttccacagttccacaatatcgcaaa 1215
 Db 1155 gacacaaatgaacgatcatcggttcatcggtttccacagttccacaatatcgcaaa 1214
 QY 1216 ctgtgtctccacacgtgcacaaatcgatcagacgcttctctcgcactcgtgtgc 1275
 Db 1215 ctgtgtctccacacgtgcacaaatcgatcagacgcttctctcgcactcgtgtgc 1274
 QY 1276 tcttggaagtcatcccggttgatgaggttggtgaagtcgcgtgtaa 1326
 Db 1275 tcttggaagtcatcccggttgatgaggttggtgaagtcgcgtgtaa 1325

RESULT 5

AAV30576
ID AAV30576 standard; DNA: 1351 BP.

AC AAV30576;

DT 07-DEC-1998 (first entry)

DE Clostridium botulinum toxin A fragment C gene in pHisBotA(syn).

KW Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;

KW botulinum; ds.

OS Clostridium botulinum serotype A.

XX Synthetic.

XX Key Location/Qualifiers

XX FT CDS 1..1338

XX FT /*tag- a

XX PN MO9808540-A1.

XX PD 05-MAR-1998.

XX PF 28-AUG-1997; 97MO-US15394.

XX PR 28-AUG-1996; 96US-0704159.

XX PA (OPHI-) OPHIDIAN PHARM INC.

XX PI Thalleys BS, Williams JA;

XX P-PI 1998-230234/20.

XX P-PSDB; AAM68391.

XX Example 29; Page 279-281; 428bp; English.

CC This is the DNA sequence of the Clostridium botulinum serotype A
 CC toxin C fragment gene contained in plasmid pHisBotA(syn). The
 CC encoded toxin A polypeptide (see AAM68391) has a histidine-tagged
 CC N-terminal extension. The vector was used to express native
 CC (i.e. non-fusion) soluble C fragment in Escherichia coli host
 CC cells. The invention relates to recombinant proteins derived from
 CC C. botulinum toxins. Methods are provided which allow for the
 CC isolation of soluble recombinant proteins free of significant
 CC endotoxin contamination. Preferred hosts for production of
 CC recombinant proteins are E. coli, insect cells and yeast cells.
 CC The recombinant toxins are used as immunogens for the production
 CC of vaccines and antitoxins that are useful in the treatment of
 CC humans and animals at risk of intoxication with clostridial toxin.
 SO Sequence 1351 BP; 407 A; 348 C; 247 G; 349 T; 0 other;

Query Match 98.4%; Score 1311; DB 19; Length 1351;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 cgtctgtctgtactctcactgaatatacaagaactatcaatctcactctgaac 75
 Db 28 cgtctgtctgtactctcactgaatatacaagaactatcaatctcactctgaac 87
 QY 76 ctgcgcgaatccaatccaatcactgaatcagctgtctcgaagttccaaatcaatc 135
 Db 88 ctgcgcgaatccaatccaatcactgaatcagctgtctcgaagttccaaatcaatc 147
 QY 136 ggttctaagttaactcgaatccgaatcgaagaatcaatccagctgtccaatc 195
 Db 148 ggttctaagttaactcgaatccgaatcgaagaatcaatccagctgtccaatc 207
 QY 196 tcttcacaaatcgaagtataccctgaagaatgctatcgatatcaactatgtagcga 255
 Db 208 tcttcacaaatcgaagtataccctgaagaatgctatcgatatcaactatgtagcga 267
 QY 256 tcttcacacccctctgtgatacccgatccgaatcactcaactcactcctgaacat 315
 Db 268 tcttcacacccctctgtgatacccgatccgaatcactcaactcactcctgaacat 327
 QY 316 gaatacacatcatalcaactgcatggaagaatctgtgtggaagaatcctgaactac 375
 Db 328 gaatacacatcatalcaactgcatggaagaatctgtgtggaagaatcctgaactac 387
 QY 376 ggtgaaatcactcgtgactcgtcgaagacactcgaagaatcaaacggtgtgttca 435
 Db 388 ggtgaaatcactcgtgactcgtcgaagacactcgaagaatcaaacggtgtgttca 447
 QY 436 tactctcagatgatacaatctctgactatacaatcgcgtgactcctgataccatc 495
 Db 448 tactctcagatgatacaatctctgactatacaatcgcgtgactcctgataccatc 507
 QY 496 aacaaatcgtgaaatcaactccaatctatcaatcaaacggtgtgttcaaacg 555
 Db 508 aacaaatcgtgaaatcaactccaatctatcaatcaaacggtgtgttcaaacg 567
 QY 556 atctccaatctgtgttaacatccacgctcttaataatcaatcgttcaaacgtgag 615
 Db 568 atctccaatctgtgttaacatccacgctcttaataatcaatcgttcaaacgtgag 627
 QY 616 cgtgacacacacgctacatctgtgatacaatcactcactcgtgtcgaagaactaac 675
 Db 628 cgtgacacacacgctacatctgtgatacaatcactcactcgtgtcgaagaactaac 687
 QY 676 gaaagaagaatcaagaacgtgtgaacacacggtccaatctgtgtatccgtgaagactc 735
 Db 688 gaaagaagaatcaagaacgtgtgaacacacggtccaatctgtgtatccgtgaagactc 747
 QY 736 tgggtgtactactcgtcgaatcgaagaacggtactcaatctgtgtatccgtgaagactc 795
 Db 748 tgggtgtactactcgtcgaatcgaagaacggtactcaatctgtgtatccgtgaagactc 807
 QY 796 aaatcgttgaacgtcaacatgtagtataccggtgttcaatcgtgaagactcgtcgtc 855
 Db 808 aaatcgttgaacgtcaacatgtagtataccggtgttcaatcgtgaagactcgtcgtc 867
 QY 856 ggttcgttgaacgtcaacatcactgaactcgtcgttcaatcgtgaagactcgtcgtc 915
 Db 868 ggttcgttgaacgtcaacatcactgaactcgtcgttcaatcgtgaagactcgtcgtc 927
 QY 916 atcatcaagaatcagctgtgtgaagaagaagaatcgttgcgaagaatcgtcgtgta 975
 Db 928 atcatcaagaatcagctgtgtgaagaagaagaatcgttgcgaagaatcgtcgtgta 987
 QY 976 tacatcaatgttgaagtaagaagaagaatcagctgtgtcgaagaatcgtcgtcgtc 1035
 Db 988 tacatcaatgttgaagtaagaagaagaatcagctgtgtcgaagaatcgtcgtcgtc 1047

QY 1036 ggtgtagaaaagatctgtctgtctgtgaaatcccgagcgttggtaactgtctcagta 1095
 |||||||
 Db 1048 ggtgtagaaaagatctgtctgtctgtgaaatcccgagcgttggtaactgtctcagta 1107
 |||||||
 QY 1096 gttgttaatgaatccaagaagcagggatcactaactaactgaatgaatgtcag 1155
 |||||||
 Db 1108 gttgttaatgaatccaagaagcagggatcactaactaactgaatgaatgtcag 1167
 |||||||
 QY 1156 gacacaatggttaacgatatcgtttcattcgtttccaccattcaacatatcgtaa 1215
 |||||||
 Db 1168 gacacaatggttaacgatatcgtttcattcgtttccaccattcaacatatcgtaa 1227
 |||||||
 QY 1216 ctgtgtcttccaactggtacatcgttcagatcgaacgttctctcgcactcgtgtgc 1275
 |||||||
 Db 1228 ctgtgtcttccaactggtacatcgttcagatcgaacgttctctcgcactcgtgtgc 1287
 |||||||
 QY 1276 tcttggagttcattcccggttgaatgaggttgggtgaacgttcgctgtaa 1326
 |||||||
 Db 1288 tcttggagttcattcccggttgaatgaggttgggtgaacgttcgctgtaa 1338
 |||||||

RESULT 6

AAT29246
 ID AAT29246 standard; DNA; 1402 BP.

AC AAT29246;

DT 07-JUL-1996 (first entry)

DE Type A neurotoxin C fragment-polyhistidine tag gene fusion.

KW Toxin; neurotoxin; fusion protein; antitoxin; vaccine; immunogen;
 Clostridium botulinum; polyhistidine; vector; PETHISA; pHISBot; ds.

OS Synthetic.

FX Key Location/Qualifiers

FT CDS 1..1317

FT /*tag= a

FT /product= pHISBot fusion protein

PN W09612802-A1.

PD 02-MAY-1996.

PF 23-OCT-1995; 95WO-US13737.

PR 07-JUN-1995; 95US-0480604.

PR 24-OCT-1994; 94US-0329154.

PR 16-MAR-1995; 95US-0405496.

PR 14-APR-1995; 95US-0422711.

PA (OPHI-) OPHIDIAN PHARM INC.

PI Pirca JR, Kink JA, Padhye NV, Stafford DC, Thalley BS;

PI Williams JA;

DR WPI; 1996-230603/23.

DR P-PSDB; AAR95009.

XX Example 24; Page 340-342; 434pp; English.

CC A nucleotide sequence (AAT29246) present in vector pETHISA encodes
 the pHISBot fusion protein (AAR95009) comprising a polyhistidine
 affinity tag and fragment C (see also AAR95008) of the Clostridium
 botulinum type A neurotoxin. The pHISBot protein was expressed
 in *Escherichia coli* as a soluble protein and was purified by
 metal chelate affinity chromatography to obtain a product free

CC of endotoxin contamination that may be useful as an immunogen
 CC in vaccine compns.
 XX
 SQ Sequence 1402 BP; 420 A; 360 C; 260 G; 362 T; 0 other;

Query Match 98.4%; Score 1311; DB 17; Length 1402;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 cgtctgtcttcaacttccatgaatcatcaatcaatcaatcaatcaatcaatcaat 75
 |||||||
 Db 79 cgtctgtcttcaacttccatgaatcatcaatcaatcaatcaatcaatcaatcaat 138
 |||||||
 QY 76 ctgtgttaagatccaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 135
 |||||||
 Db 139 ctgtgttaagatccaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 198
 |||||||
 QY 136 gttctcaagttcaactcgtatccgacgaagaatcagatccaggtgttcaatcgtgaa 195
 |||||||
 Db 199 gttctcaagttcaactcgtatccgacgaagaatcagatccaggtgttcaatcgtgaa 258
 |||||||
 QY 196 tcttccaaaatcgaagttaactcgtgaagaatgctatcgtatcaactctatgtacgaagaac 255
 |||||||
 Db 259 tcttccaaaatcgaagttaactcgtgaagaatgctatcgtatcaactctatgtacgaagaac 318
 |||||||
 QY 256 tcttccaaactccttctggtatccgataccgaataacttcaactcctctctgaacaat 315
 |||||||
 Db 319 tcttccaaactccttctggtatccgataccgaataacttcaactcctctctgaacaat 378
 |||||||
 QY 316 gaatacacatcatcaacgcatggaagaacaaatctcgtgttgaagaatctatctgaaactac 375
 |||||||
 Db 379 gaatacacatcatcaacgcatggaagaacaaatctcgtgttgaagaatctatctgaaactac 438
 |||||||
 QY 376 ggtgaatcatctgtgactctgcaagacactcaggaatcaacaacggttctgtatcaaa 435
 |||||||
 Db 439 ggtgaatcatctgtgactctgcaagacactcaggaatcaacaacggttctgtatcaaa 498
 |||||||
 QY 436 tactctcagaatgatacaactctctgatacatcaatcgttctgttcttccatcaac 495
 |||||||
 Db 499 tactctcagaatgatacaactctctgatacatcaatcgttctgttcttccatcaac 558
 |||||||
 QY 496 aacaatcgtctgaataacccaacaaatcatatcaatcagcgttctgtatcgaaccgaagaacg 555
 |||||||
 Db 559 aacaatcgtctgaataacccaacaaatcatatcaatcagcgttctgtatcgaaccgaagaacg 618
 |||||||
 QY 556 atctccaatctggttaacaatccaagcgttcttaataacatcatgttcaactgtgaaggtgt 615
 |||||||
 Db 619 atctccaatctggttaacaatccaagcgttcttaataacatcatgttcaactgtgaaggtgt 678
 |||||||
 QY 616 cgtgacactcaccggtacactcgtgataactcaatctctgttgcagaagaagactggaac 675
 |||||||
 Db 679 cgtgacactcaccggtacactcgtgataactcaatctctgttgcagaagaagactggaac 738
 |||||||
 QY 676 gaataaagaatcaaaagactgttacgacaacacgctccaatctgtgttccctgaaagacttc 735
 |||||||
 Db 739 gaataaagaatcaaaagactgttacgacaacacgctccaatctgtgttccctgaaagacttc 798
 |||||||
 QY 736 tgggtgtactactcgtcagtaagacaacacggttactatcatgtcgaatctgtacatccgaac 795
 |||||||
 Db 799 tgggtgtactactcgtcagtaagacaacacggttactatcatgtcgaatctgtacatccgaac 858
 |||||||
 QY 796 aataagttgagcgtcaaaatgtagttatccggttgaatgaatgaatgaatgaatgaatgaat 855
 |||||||
 Db 859 aataagttgagcgtcaaaatgtagttatccggttgaatgaatgaatgaatgaatgaatgaatgaat 918
 |||||||
 QY 856 ggttctgtatgactactcaacaatctacacgtgaactctccctgtacggtgtgtacaaactc 915
 |||||||
 Db 919 ggttctgtatgactactcaacaatctacacgtgaactctccctgtacggtgtgtgtacaaactc 978
 |||||||
 QY 916 atcatcaagaataacgctctgtgttaacaaagacaatcatcgttgcgaacaatgtatcgtgtga 975
 |||||||
 Db 979 atcatcaagaataacgctctgtgttaacaaagacaatcatcgttgcgaacaatgtatcgtgtga 1038
 |||||||

QY 976 tacatcaatgtctgttagttaagaacaaataaccgtctgtgtaccgaatgtcttcaggtc 1035
 |||||||
 Db 1039 tacatcaatgtctgttagttaagaacaaataaccgtctgtgtaccgaatgtcttcaggtc 1098
 QY 1036 ggtgtatgaagaagctctgtctgtctgtgaatacccgagcttgtaactcgtctcagta 1095
 |||||||
 Db 1099 ggtgtatgaagaagctctgtctgtctgtgaatacccgagcttgtaactcgtctcagta 1158
 QY 1096 gttgtatgaataccaagaacagccaggtgtatctacttaacaaatgtaaatcttcag 1155
 |||||||
 Db 1159 gttgtatgaataccaagaacagccaggtgtatctacttaacaaatgtaaatcttcag 1218
 QY 1156 gacaacaatgtgtaacgatatcgtgttccatcgtttccaccagttccaacaaatcgcctaa 1215
 |||||||
 Db 1219 gacaacaatgtgtaacgatatcgtgttccatcgtttccaccagttccaacaaatcgcctaa 1278
 QY 1216 ctgtgtgtcttcccaactcgtgtaacatcgttcgaatcgttcctcctcgtcactcgtgttc 1275
 |||||||
 Db 1279 ctgtgtgtcttcccaactcgtgtaacatcgttcgaatcgttcctcctcgtcactcgtgttc 1338
 QY 1276 tcttggagatcaccggtgtgtaacggttgggtgaacgctcgcctgttaa 1326
 |||||||
 Db 1339 tcttggagatcaccggtgtgtaacggttgggtgaacgctcgcctgttaa 1389

RESULT 7
 AAV30572
 ID AAV30572 standard; DNA; 1402 BP.

AAV30572;

07-DEC-1998 (first entry)

Clostridium botulinum toxin A fragment C gene in phisBot.

Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;

botulinum; ds.

Clostridium botulinum serotype A.

Synthetic.

Key CDS location/Qualifiers

FT CDS 1..1389

/+tag= a

W09808540-A1.

05-MAR-1998.

28-AUG-1997; 97WO-US15394.

28-AUG-1996; 96US-0704159.

(OPHI-) OPHIDIAN PHARM INC.

Thalley BS, Williams JA;

WPI; 1998-230234/20.

P-PSDB; AAW68390.

Host cell containing recombinant expression vector encoding Clostridium botulinum type B or E toxin - useful to treat humans and other animals at risk of intoxication with clostridial toxin

Example 24; Page 265-267; 428pp; English.

This is the DNA sequence of the Clostridium botulinum serotype A toxin C fragment gene contained in plasmid phisBot. The encoded toxin A polypeptide (see AAW68390) has a histidine-tagged N-terminal extension. The vector was used to express native (i.e. non-fusion) soluble C fragment in Escherichia coli host cells. The invention relates to recombinant proteins derived from

CC C. botulinum toxins. Methods are provided which allow for the
 CC isolation of soluble recombinant proteins free of significant
 CC endotoxin contamination. Preferred hosts for production of
 CC recombinant proteins are E. coli, insect cells and yeast cells.
 CC The recombinant toxins are used as immunogens for the production
 CC of vaccines and antitoxins that are useful in the treatment of
 CC humans and animals at risk of intoxication with clostridial toxin.
 XX

Sequence 1402 BP; 420 A; 360 C; 260 G; 362 T; 0 other;

Query Match 98.4%; Score 1311; DB 19; Length 1402;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 cgtctgtgtctcctcctcctcctgaataataagaacatcaataactcctcctgaac 75
 |||||||
 Db 79 cgtctgtgtctcctcctcctcctgaataataagaacatcaataactcctcctgaac 138
 QY 76 ctgcgtacgaataccaataccaactcgtatcgtctcgtcagcttccaataccaatc 135
 |||||||
 Db 139 ctgcgtacgaataccaataccaactcgtatcgtctcgtcagcttccaataccaatc 198
 QY 136 ggtcttaagttcaacttcgtatccgatccgaagaatcagatccaagctgttcaatctgtaa 195
 |||||||
 Db 199 ggtcttaagttcaacttcgtatccgatccgaagaatcagatccaagctgttcaatctgtaa 258
 QY 196 tcttccaagaatcgaatgtatccctcgaagaatcgtcgtatcacaactcgtatcgaagaac 255
 |||||||
 Db 259 tcttccaagaatcgaatgtatccctcgaagaatcgtcgtatcacaactcgtatcgaagaac 318
 QY 256 tcttccaactcctctcgtgtaacgtatcccgaaatcacttcaactcctcctcgaacat 315
 |||||||
 Db 319 tcttccaactcctctcgtgtaacgtatcccgaaatcacttcaactcctcctcgaacat 378
 QY 316 gaataccaatcaataactcgtatcgtgtaagaatcgttgcgaactc 375
 |||||||
 Db 379 gaataccaatcaataactcgtatcgtgtaagaatcgttgcgaactc 438
 QY 376 ggtgaatactcgtgtaactcgtcgaagacactcgaagaatcacaacgctgtgtatcaaa 435
 |||||||
 Db 439 ggtgaatactcgtgtaactcgtcgaagacactcgaagaatcacaacgctgtgtatcaaa 498
 QY 436 tactctcaagatgaacaactcgttgaactacaataatcgtgtgacttaccatcacc 495
 |||||||
 Db 499 tactctcaagatgaacaactcgttgaactacaataatcgtgtgacttaccatcacc 558
 QY 496 aagaatcgttgaataactcacaatactatacaacgctgtgtgactcgaagaacg 555
 |||||||
 Db 559 aagaatcgttgaataactcacaatactatacaacgctgtgtgactcgaagaacg 618
 QY 556 atctccaatctggtgaacaacacgcttcaatacaactcgttcaaacgtgagctgtg 615
 |||||||
 Db 619 atctccaatctggtgaacaacacgcttcaatacaactcgttcaaacgtgagctgtg 678
 QY 616 cgtgacactcaccgtctacatctgtaacaaatctcaatctgttctgcacaagaactgaac 675
 |||||||
 Db 679 cgtgacactcaccgtctacatctgtaacaaatctcaatctgttctgcacaagaactgaac 738
 QY 676 gaataagaatacaagaacgtgtacgaacaacatccaatctgtgttccgaaagcttc 735
 |||||||
 Db 739 gaataagaatacaagaacgtgtacgaacaacatccaatctgtgttccgaaagcttc 798
 QY 736 tgggtgtgactaccgtgtaacgaacaacgttactatcgttgaatcgttgcgtaccgaac 795
 |||||||
 Db 799 tgggtgtgactaccgtgtaacgaacaacgttactatcgttgaatcgttgcgtaccgaac 858
 QY 796 aatacgttgaacgtcaacaatgtagtataccgcggttacaatgtaactgaaagctcgct 855
 |||||||
 Db 859 aatacgttgaacgtcaacaatgtagtataccgcggttacaatgtaactgaaagctcgct 918
 QY 856 ggtctgtatgactccaacatctcactcgaactctcctcgtgacggtgtaccatc 915
 |||||||

XX	New nucleic acids encoding the carboxy- or amino-terminal portions of
PT	the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT	vaccine against botulism
xx	
PS	Disclosure: Fig 2a; 73pp; English.
XX	
CC	Botulinm neurotoxins are translated as a single 150 kDa polypeptide
CC	chain and then posttranslationally nicked, forming a dichain
CC	consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC	remain linked by a disulfide bond. Nucleic acids encoding the
CC	carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC	chain of Botulinum neurotoxin (BoNT) can be used in recombinant
CC	expression vectors and expressed in transformed cells to produce
CC	peptide antigens useful for eliciting an immune response to give
CC	protective immunity against botulinum neurotoxin, which causes
CC	botulism. The nucleic acids are expressible in a recombinant
CC	organisms such as Escherichia coli or Pichia pastoris. The use
CC	of recombinant nucleic acids are advantageous since it eliminates
CC	the need to culture large quantities of hazardous toxin-producing
CC	bacterium. Production yield from the genetically engineered product
CC	is also high and cost of production is lower. The nucleic acids can
CC	be derived from Clostridium botulinum serotypes A-G.
xx	
SQ	Sequence 1323 BP; 404 A; 334 C; 242 G; 343 T; 0 other;
Query Match	98.4%; Score 1310.4; DB 22; Length 1323;
Best Local Similarity	99.9%; Pred.NO. 0;

OY	21	gctgtctaccttcaactgtaatatacaaagaacatcatcaatcaccaccaaccctgaaacctgcg	80
OY	12	gattgtacacttccatcagtaatacatcaagaacatacacaatcattccattccattgaaacctgcg	71
OY	81	ctacgaatccaatacaaccctgatactgacccctgtctgcctacagcttccaaatcaaatcggttc	140
Db	72	ctacgaatccaatacaaccctgatactgacccctgtctgcctacagcttccaaatcaaatcggttc	131
OY	141	taaaagttaactcgatccgatctgcagacaagaatcagatccagctgttccaatctggaatcttc	200
Db	133	taaaagttaactcgatccgatctgcagacaagaatcagatccagctgttccaatctggaatcttc	191
OY	201	caaaatcgagaagtatccctcgaagaatgctatcgttatcaaaccttaigttaagaanaacttcc	260
Db	192	caaaatcgagaagtatccctcgaagaatgctatcgttatcaaaccttaigttaagaanaacttcc	251
OY	261	caaccctctctgtaactcgtatcccggaataactccaactcattccctctgtaagaatgata	320
Db	252	caaccctctctgtaactcgtatcccggaataactccaactcattccctctgtaagaatgata	311
OY	321	caccatctatcaactgcatactggaanaacaattcttggttggaaagtatctctgtaactacggtga	380
Db	312	caccatctatcaactgcatactggaanaacaattcttggttggaaagtatctctgtaactacggtga	371
OY	381	aatcatctggaactcgcgaagacacccaggaataccaacacgcgtgtgtgatactcaaatatcc	440
Db	372	aatcatctggaactcgcgaagacacccaggaataccaacacgcgtgtgtgatactcaaatatcc	431
OY	441	tcaagatgatacaacatctcttgactacaatcaactcgtctgagatcttcgttatccataccaacaa	500
Db	432	tcaagatgatacaacatctcttgactacaatcaactcgtctgagatcttcgttatccataccaacaa	491
OY	501	tctgtctgaaatactccaanaattacataccaagcgccgtctgtaactcgagacagaacccgattcc	560
Db	492	tctgtctgaaatactccaanaattacataccaagcgccgtctgtaactcgagacagaacccgattcc	551
OY	561	caactcgggttaacatccacagccttctaataacaatcabytccaactggaacggttgcgtga	620
Db	552	caactcgggttaacatccacagccttctaataacaatcabytccaactggaacggttgcgtga	611
OY	621	caactacccctcatctgatactcaaatataacttcaatctggttcgacaagaactggaacgaaa	680

Db	612	caaccacgctacacatctcgtatccaatacttcaatctctgttcgacaaagaactgacgaanaa	671
Qy	681	agaaatacaagaacctgtacgacaaaccacgctccaattctcgtatccctgaagaactctcggg	740
Db	672	agaaatacaagaacctgtacgacaaaccacgctccaattctcgtatccctgaagaactctcggg	731
Qy	741	tgactacctgtgacgacgaacaaccgctactacatgacctgaactctgacatccgacaataa	800
Db	732	tgactacctgtgacgacgaacaaccgctactacatgacctgaactctgacatccgacaataa	791
Qy	801	cgttgcacgtcaacaatgtaaggtatcccggtttacaatgtatccctgaagaagtcgcgctgttc	860
Db	792	cgttgcacgtcaacaatgtaaggtatcccggtttacaatgtatccctgaagaagtcgcgctgttc	851
Qy	861	tgattatgactccaacgtcttaccctggaactcttccctcgtctaccgcggtgacccaattctaat	920
Db	852	tgattatgactccaacgtcttaccctggaactcttccctcgtctaccgcggtgacccaattctaat	911
Qy	921	caagaataacgcgctctcgtgtlaagacaagacaatatcgtttcgcgaacaactgatatcat	980
Db	912	caagaataacgcgctctcgtgtlaagacaagacaatatcgtttcgcgaacaactgatatcat	971
Qy	981	caatgctgttagttaagaacaagaataacgcgtctcgtgtactacaatgcttctaaagctcgtct	104
Db	972	caatgctgttagttaagaacaagaataacgcgtctcgtgtactacaatgcttctaaagctcgtct	103
Qy	1041	agaaaaagatctctgcctcgtcgtgaatacccggaacgcgtgtgtaactcgtctcaagttagt	110
Db	1032	agaaaaagatctctgcctcgtcgtgaatacccggaacgcgtgtgtaactcgtctcaagttagt	109
Qy	1101	aatgaaatccaagaagaacgacagggtatcactaacaatactgcaaaatgatatctgcagacaa	116
Db	1092	aatgaaatccaagaagaacgacagggtatcactaacaatactgcaaaatgatatctgcagacaa	115
Qy	1161	caatggttaacagatatcggtttcatccggtttccaccacagttccaacaatacgtctaaactggt	122
Db	1152	caatggttaacagatatcggtttcatccggtttccaccacagttccaacaatacgtctaaactggt	121
Qy	1221	tgcttccaactgtagtaacaactcgttcagatccggaacggttccctcgcgactctgggttgctcttg	128
Db	1212	tgcttccaactgtagtaacaactcgttcagatccggaacggttccctcgcgactctgggttgctcttg	127
Qy	1281	ggaaattatcccggttgtatgacggttggtgggtgaaagcgcgcgtgtaagaattc	1332
Db	1272	ggaaattatcccggttgtatgacggttggtgggtgaaagcgcgcgtgtaagaattc	1323

```

RESULT      9
AAA54484
ID   AAA54484 standard; DNA; 1326 BP.

```

DT	11-APR-2001	(first entry)
XX		
DE	botulism toxin heavy chain C-terminal coding sequence (serotype A).	
XX		
KW	Botulism; toxin; neurotoxin; heavy chain; recombinant expression;	
KW	recombinant vector; antigen; immune response; vaccine; bacterium;	
KM	infection; ds.	
XX		
OS	Synthetic.	
OS	Clostridium botulinum.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	13..1320
FT		/*tag a
FT		/product=H_C peptide fragment
XX		
PN	MO200067700-A2.	
XX		
PD	16-NOV-2000.	
XX		

```

PF 12-MAY-2000; 2000OMO-US12890.
XX 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.
PR 12-MAY-1999; 99US-0133869.
PR 12-MAY-1999; 99US-0133873.
PR 29-JUL-1999; 99US-0146192.
PA (USSA ) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
XX Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
PI WPI: 2001-016048/02.
DR P-PSDB: AAB04090.
XX
XX New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulism
XX
XX Disclosure: Fig 3a; 73pp; English.
XX
XX Botulism neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dicanin
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.
XX
XX Sequence 1326 BP; 404 A; 336 C; 243 G; 343 T; 0 other;
SQ
Query Match 98.2%; Score 1308.2; DB 22; Length 1326;
Best Local Similarity 99.8%; Pred. NO. 0;
Matches 1310; Conservative 0; Mismatches 3; Indels 0; Gaps
0;
OY 20 tgcgtctacccttaccatatacatcaagaacatcatcaatcctccatcctgaacctgc 79
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 14 tgcctctacccttaccatatacatcaagaacatcatcaatcctccatcctgaacctgc 73
OY 80 gctacgaatccaatactatcgatcgacgcgctctcgtcgtacgctccaataatcaacatcggt 139
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 74 gctacgaatccaatactatcgatcgacgcgctctcgtcgtacgctccaataatcaacatcggt 133
OY 140 ctaaggttaactcgcatacgcgcgaagaatcagatccagtcggttcacatcgcgaatcct 199
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 134 ctaaggttaactcgcatacgcgcgaagaatcagatccagtcggttcacatcgcgaatcct 193
OY 200 ccaaaatcgaaatatacctgaagaatgctatcgtatacaactatgtiaagaaactct 259
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 194 ccaaaatcgaaatatacctgaagaatgctatcgtatacaactatgtiaagaaactct 253
OY 260 ccaactctcttgcggtatccgataccgaaatcttaactccatctcctcgaacaatgat 319
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 254 ccaactctcttgcggtatccgataccgaaatcttaactccatctcctcgaacaatgat 313
OY 320 acaccatcatcaacgcgatgtaaaacaattctcgtgttggaagatatactctgaactacggtg 379
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 314 acaccatcatcaacgcgatgtaaaacaattctcgtgttggaagatatactctgaactacggtg 373
OY 380 aatcatatggaactctgcaagacactcaggaataaacaagcgtgttctatcaaatct 439
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||

```

Sequence 1326 BP; 404 A; 336 C; 243 G; 343 T; 0 other;

Key Match	98.2%	Score 1308.2	DB 22	Length 1326
st Local Similarity	99.8%	Pred No. 0		
ches 1310	Conservative	0	Mismatches 3	Indels 0
				Gaps 0
20	tgtgtctacatcctactgaatacatcatcaagaacatcatcaatactccatccctgaacctgc	79		
1				
14	tggccctacatcctactgaatacatcatcaagaacatcatcaatactccatccctgaacctgc	73		
80	gctagaatccaatcgaactgatacgcctgtcttcgtctaagcttcgaaatcaaatcatcggtt	139		
74	gtacgaatccaatcgaactgatacgcctgtcttcgtctaagcttcgaaatcaaatcatcggtt	133		
140	ctaaagttaacttcggttcgatccgagaagaatcagttccagttgttcaatctcgaaatcct	199		
134	ctaaagttaacttcggttcgatccgagaagaatcagttccagttgttcaatctcgaaatcct	193		
200	ccaagaatcgaagtctatccggaagaatcgcgtatcgatatacaactctatgtatcgaaaccttc	259		
194	ccaaatcgaagtctatccggaagaatcgcgtatcgatatacaactctatgtatcgaaaccttc	253		
260	ccacccctctcttgatccglatccggaatcactcgaactcgaactcctctcttgaaacaatgaa	319		
254	ccacccctctcttgatccglatccggaatcactcgaactcgaactcctctcttgaaacaatgaa	313		
320	acaaccatcatcaactcgcgcgaaacaactctcgttctggaaglatccctgaaactcgcgtg	379		
314	acaaccatcatcaactcgcgcgaaacaactctcgttctggaaglatccctgaaactcgcgtg	373		
380	aaatcatctggaactcgtcgaagaactcgaagaatcaaacgcgtgttattcaaatcct	439		

Db 374 aaatcatctgactctgacggaacatcaaggaaatcaacsgcgtgtgtatctcaact 433
 QY 440 ctcaatgatcaacatctctgactacatcatcgctgactcttctgtatcatccaca 499
 Db 434 ctcaatgatcaacatctctgactacatcatcgctgactcttctgtatcatccaca 493
 QY 500 atcgtctgaatactccaataatctacatcaacggtcgctgtatcgacagaacgact 559
 Db 494 atcgtctgaatactccaataatctacatcaacggtcgctgtatcgacagaacgact 553
 QY 560 ccaatctgggttaacatccacgcttctataatcatcagtccaactggaagtggtg 619
 Db 554 ccaatctgggttaacatccacgcttctataatcatcagtccaactggaagtggtg 613
 QY 620 acactcaacgctcatctgtgatcaataatctctgttcgcgaagaagaactggaagaa 679
 Db 614 acactcaacgctcatctgtgatcaataatctctgttcgcgaagaagaactggaagaa 673
 QY 680 aagaatcaagaacctgttaacgaacacagtcacaattctgttlatctgaagaactctgg 739
 Db 674 aagaatcaagaacctgttaacgaacacagtcacaattctgttlatctgaagaactctgg 733
 QY 740 gtaactactctgacgacgaacacgctactacatgtctgaatctgtacgactcgaacaat 799
 Db 734 gtaactactctgacgacgaacacgctactacatgtctgaatctgtacgactcgaacaat 793
 QY 800 acgttgcgtcaacaatgtagtgcgtgcgtgtacatgtagtgcgtgcgtgcgtgcgtg 859
 Db 794 acgttgcgtcaacaatgtagtgcgtgcgtgtacatgtagtgcgtgcgtgcgtgcgtg 853
 QY 860 ctgttactgaacccaacatctactgtaactcttccctgtacacgctgcgtgcgaacatctaca 919
 Db 854 ctgttactgaacccaacatctactgtaactcttccctgtacacgctgcgtgcgaacatctaca 913
 QY 920 tcaagaatacgcgtctgcgtgaacagaacatcatcgttcgaacaatgactcgtgtatca 979
 Db 914 tcaagaatacgcgtctgcgtgaacagaacatcatcgttcgaacaatgactcgtgtatca 973
 QY 980 tcaatgtgtgagttaagaacaaagaatctacgctgtgcgttaccatgcttccagcgtgtg 1039
 Db 974 tcaatgtgtgagttaagaacaaagaatctacgctgtgcgttaccatgcttccagcgtgtg 1033
 QY 1040 tagaanaagatctgtctgtctctggaatcccggaacgttggtaatctgtcgaagtagtg 1099
 Db 1034 tagaanaagatctgtctgtctctggaatcccggaacgttggtaatctgtcgaagtagtg 1093
 QY 1100 taatgaataccaagaagaccaggtgtatcaactaacaatgcaaaatgaaatctgcagaga 1159
 Db 1094 taatgaataccaagaagaccaggtgtatcaactaacaatgcaaaatgaaatctgcagaga 1153
 QY 1160 acaatgttaacgatatcggttctacgttccacaggttccaaatcatcgaatactg 1219
 Db 1154 acaatgttaacgatatcggttctacgttccacaggttccaaatcatcgaatactg 1213
 QY 1220 ttgcttcaacactggtataactcgttcagaatcgaaacttctcctgcgactcgtgtgtctt 1279
 Db 1214 ttgcttcaacactggtataactcgttcagaatcgaaacttctcctgcgactcgtgtgtctt 1273
 QY 1280 gggagttcatcccggtgtgtagagcgtgtgtgtgaacgttcgcgtgtgaagaattc 1332
 Db 1274 gggagttcatcccggtgtgtagagcgtgtgtgtgaacgttcgcgtgtgaagaattc 1326
 RESULT 10
 AAA54588
 ID AAA54588 standard; DNA: 1332 BP.
 AC AAA54588:
 XX
 XX 11-APR-2001 (first entry)
 DT
 DE Sequence encoding botulism toxin C fragment (serotype A).
 XX

KW Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
 KM recombinant vector; antigen; immune response; vaccine; bacterium;
 KM Infection; ds.
 XX
 OS Synthetic.
 OS Clostridium botulinum.
 PN WO200067700-A2.
 XX
 PD 16-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US12890.
 XX
 PR 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.
 XX
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX
 PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 PI
 XX
 DR WPI; 2001-016048/02.
 XX
 PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism
 PS
 PS
 XX
 XX Example 7; Page 37; 73pp; English.
 CC Botulism neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a di-chain
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.
 XX
 SQ Sequence 1332 BP; 401 A; 342 C; 244 G; 344 T; 1 other;
 Query Match 94.3%; Score 1256.6; DB 22; Length 1332;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1303; Conservative 1; Mismatches 1; Indels 6; Gaps 4;
 QY 16 cgtctgtgtctactcactgaatacatcaagaacatcatcactcatctgaac 75
 Db 15 cgtctgtgtctactcactgaatacatcaagaacatcatcactcatctgaac 74
 QY 76 ctgcgtacgaatcaatcaactatcgactcgtctcgtacgttccaaatcaaatc 135
 Db 75 ctgcgtacgaatcaatcaactatcgactcgtctcgtacgttccaaatcaaatc 134
 QY 136 ggtctaaagttaactcgcgatcgcagacaagaatcagatccagctgttcaatcggaa 195
 Db 135 ggtctaaagttaactcgcgatcgcagacaagaatcagatccagctgttcaatcggaa 194
 QY 196 tcttcaaaaatcgaagtatcttgaagaatgctatcgtatcaactatgtagaagaac 255
 Db 195 tcttcaaaaatcgaagtatcttgaagaatgctatcgtatcaactatgtagaagaac 254

OY	286	ttccacgctctctcttgatcgcgtatcccgaaattacttaactcactcctctgaaat	315
Db	285	ttctccacctctctcttgatctcgatctcc-aaatacttaactccaactcctctgaaat	313
OY	316	gaatacaccaatcaacaactgcgatgaaanaaactctcgtctggaagtaactctgaaac	375
Db	314	gaatacaccaatcaatcactgcgatgaaanaaactctcgtctggaagtaactctgaaac	373
OY	376	ggtgaatacatctggaactctgcagagacatccaggaatacaaacgcgtctgtattcaa	435
Db	374	ggtgaatacatctggaactctgcagagacatccaggaatacaaacgcgtctgtattcaa	433
OY	436	tactctcaagatgaatcaactctcgtacataatcgcgtgcgtgcgttcgttaactaac	495
Db	434	tactctcaagatgaatcaactctcgtacataatcgcgtgcgtgcgttcgttaactaac	493
OY	496	aacaaatcgctgtaataactccaaaatctacatacaacgcgcgtctgatacgacaagacg	555
Db	494	aacaaatcgctgtaataactccaaaatctacatacaac-gcgcgtctgatacgacaagacg	552
OY	556	atctccaaatcgtggataacatccacgcgtctcctaataataatggttaaacatgaaagttg	615
Db	553	atctccaaatcgtggataacatccacgcgtctcctaataataatggttaaacatgaaagttg	611
OY	616	cgtgcacatccacgcgtacatctcgtatcaaatctgttcgaacaaagaactgac	675
Db	612	cgtgcacatccacgcgtacatctcgtatcaaatctgttcgaacaaagaactgac	671
OY	676	gaaaaaagaatacaaaagccgttaacgaaacaaatccaattctggtatccctgaaagattc	735
Db	672	gaaaaaagaatacaaaagccgttaacgaaacaaatccaattctggtatccctgaaagattc	731
OY	736	tgggtgtgactacccgcgcagtaacgaacacgcgtactacatgctgtaactgtacatccgaac	795
Db	732	tgggtgtgactacccgcgcagtaacgaacacgcgtactacatgctgtaactgtacatccgaac	791
OY	796	aaatacgttgaacgtccaacaatgtatgatatccgcggttaacatgtactgaaaggtccgcgt	855
Db	792	aaatacgttgaacgtccaacaatgtatgatatccgcggttaacatgtactgaaaggtccgcgt	851
OY	856	ggtctgtgatactacacaacatctaacctgaactctccctgttaacggtgtgatacaaatc	915
Db	852	ggtctgtgatactacacaacatctaacctgaactctccctgttaacggtgtgatacaaatc	911
OY	916	atcatcaagaataacgcgcgtcgtgtaacaaaggaaataatcgttcgacaacaatgactgta	975
Db	912	atcatcaagaataacgcgcgtcgtgtaacaaaggaaataatcgttcgacaacaatgactgta	968
OY	976	tactatcaatgttgaatgaatgaanaaagaatataccggtctggtctacaaatgcttccaagt	1033
Db	968	tactatcaatgttgaatgaatgaanaaagaatataccggtctggtctacaaatgcttccaagt	1028
OY	1036	ggtgtgagaagaagactcgtcgtcgtcgtgaaatcccggaacggttgtaactgtctcaagta	1095
Db	1029	ggtgtgagaagaagactcgtcgtcgtcgtgaaatcccggaacggttgtaactgtctcaagta	1088
OY	1096	gtctgtaatgaatacccaagaacgcagcaggtatcaactaaacaaatgcaaaatgaaatcgtgac	1155
Db	1089	gtctgtaatgaatacccaagaacgcagcaggtatcaactaaacaaatgcaaaatgaaatcgtgac	1148
OY	1156	gacacaaatgtataacgatacgtgttatacgtgttccacaaagttaacaatatgctgttaa	1215
Db	1149	gacacaaatgtataacgatacgtgttatacgtgttccacaaagttaacaatatgctgttaa	1206
OY	1216	ctggttgccttccaactgtgatacaatcgtcaaatcgaatcgaatctccctcgtcgaactctggtgc	1277
Db	1209	ctggttgccttccaactgtgatacaatcgtcaaatcgaatcgaatctccctcgtcgaactctggtgc	1266
OY	1276	tcttggaaggttcaatcccggtgtgatacgcgtgttggtgtgaacgtccgcgtgtaa	1326
Db	1269	tcttggaaggttcaatcccggtgtgatacgcgtgttggtgtgaacgtccgcgtgtaa	1319

QY 203 aaatcgaagttatccctggaagaatgctacgtatatacaactctatgtaacgaaactctcca 262
 Db 373 aaatgagtgtaatttttaaaaaatgctatgtatataatagatagatgataaaattttatga 432
 QY 263 cctcctctggtatccgtaaccgaatacttcaactcctcctctgtaacaatgataca 322
 Db 433 ctagtcttctgtaagaatactcctaagatatttaacagatataagcttaataatgataata 492
 QY 323 ccatcataactgctatgtaaaaaaactctgtgtggaagtatctctgtaactaggtgtaa 382
 Db 493 caataataaattgtatggaataaataatcaggaatggaagaatcacttaattatggtgtaa 552
 QY 383 tcatctgactctgcaagcactcaggaatacaaacagcggtgtgtattcaataactctcc 442
 Db 553 taactctgacttatacagataactcaggaataaacaagaatgatttttaataacagtc 612
 QY 443 agatgatcaacatctctgtaactacatcaatcgtgtgactctgttaccatcaacaacatc 502
 Db 613 aaatgattatataatcagattatataaacaagatgatttttgtaactactcaataata 672
 QY 503 gtcgtgaataactccaaatctacatcaacggtctgtatgtaacacgaacggaatctcca 562
 Db 673 gattaaataaacttaaaatctatatacaaatggaagaatagatcaaaaacaaacttcaaa 732
 QY 563 atctggtgaacatccacgcttcaataacatcatgattcaaacgtgagctgtctgtgaca 622
 Db 723 atttaggtaataatcatctgctagtaataatagtttaataatgattggtgtgtaagata 792
 QY 623 ctacacgtacatctgtaataactcaatctcaatctgtctgcacaagaactgtaagaagaa 682
 Db 793 cactagatataattgtgaataaataatttaactcttttgatagaagaataatgaagaaag 852
 QY 683 aaataaagaacctgtagcaaacacgcttcaactctggtatctgtaacgaactcttggtgtg 742
 Db 853 aaataaagaatttataatgataatcaacaatcaggtatctttaaagaactcttggtgtg 912
 QY 743 actacgtgacgtacgacaacacgttactacatgctgtacatctgcgaacaataacg 802
 Db 913 attattacaataatgataaacaactactatattgttaatttataatgataccaataatg 972
 QY 803 ttgacgtcaacaatgtaaggtatcggttataatgtaacgtgaaggtcggtgtctgtg 862
 Db 973 tctgattgtaataatgtaaggtatgaggttataatgtaacgtgaaggtcaggtgtaag 1032
 QY 863 ttatgactacacacatctgtaactctctccctgtacccgtggtgaacaaattcatca 922
 Db 1033 taatgactacacacatcttatttataatcaagttgtatagggggaacaaatttattataa 1092
 QY 923 agaaatacgtctgtgtaacaagaacatatacgtctgcgaacaatgctgtatcatca 982
 Db 1093 aaaaatagctctggaataaagaataatattgttagaataatgataatgctgtatcat 1152
 QY 983 atgtgtgtgtaagaacaagaataacgtctgtgtaacaaatgctctcaggtgtgtaag 1042
 Db 1153 atgtgtgtgttaaaaaataaagaataataggttagcttaacttaacacaggaagcgtaag 1212
 QY 1043 aaagatctgtctgtctgtggaatacccggaacgttggtaactgtctcaggtgtaag 1102
 Db 1213 aaaaatctactagctatgataaataactgtatgtaggaataatcuaagtaagaagtaaa 1272
 QY 1103 tgaataccaagaacgacaggtgtatcactaaacaaatgcaaaatgtaactgcaagcaaca 1162
 Db 1273 tgaagtcataaaaatgataagaagaatacaaaatgaatgcaaaatgtaactcaagtaata 1332
 QY 1163 atgtgtaagatagctgttcatcgttccacacagttcaacaataatcgtctaaactgtgtg 1222
 Db 1333 atggaatgataatagctcttataagattcatcagtttaataataatcaaaactagtag 1392
 QY 1223 ctccaactgtgtacaaatgctgaagatcgaacgttctcgcacactggtgtgcttctggg 1282
 Db 1393 caagttaattgttataatagacaaataagaagaatcagtaggaacttgggtgtgtcattg 1452

QY 1283 agttacccggttgatgacggttggtggaacgtccgctgtgaagaatctc 1332
 Db 1453 aattatctcgtgtagatggtggtggaagaagccactgtgaataatc 1502

RESULT 12
 AAC64582
 ID AAC64582 standard; cDNA: 4835 BP.
 XX
 AAC64582;
 AC
 XX
 DT 15-FEB-2001 (first entry)
 DE
 XX BOTM/A neurotoxin prototoxin nucleotide sequence SEQ ID NO:10.
 KW Human; procholecystokinin; CCK A receptor; CCK B receptor;
 XX pancreatitis; antiinflammatory; ss.
 OS Clostridium botulinum.
 XX
 PN WO200061192-A2.
 PD 19-OCT-2000.
 PF 06-APR-2000; 2000WO-US09142.
 PR 08-APR-1999; 99US-0288326.
 PA (ALLR) ALLEGAN SALES INC.
 PI Steward LE, Sachs G, Aoki KR;
 XX
 DR WPI: 2000-679416/66.
 PT New composition for treating acute pancreatitis, comprises a pancreatic
 PT cell surface marker binding element, a translocation element that
 PT transfers polypeptide across vesicular membrane, and a therapeutic
 PT element -
 PS Disclosure: Page 29-32; 50pp; English.
 CC The present invention describes a composition (I) for treating acute
 CC pancreatitis. (I) comprises a first element containing a binding element
 CC that binds to a pancreatic cell surface marker, a second element
 CC containing a translocation element that facilitates polypeptide transfer
 CC across the vesicular membrane, and a third element containing a
 CC therapeutic element that inhibits enzyme secretion in pancreatic cell
 CC cytoplasm. Also described is a method for making a therapeutic
 CC polypeptide having a binding element selective for cholecystokinin (CCK)
 CC receptor by expressing within a host cell a recombinant chimeric
 CC polypeptide comprising an extein containing a therapeutic element and a
 CC translocational element, and an intein located to the carboxy terminal
 CC of extein having a cysteine, serine or threonine at its amino terminus,
 CC and contacting the extein with a synthetic peptide comprising a CCK
 CC amino acid sequence containing an amidated phenylalanine at a natural
 CC C-terminus, and a cysteine, serine or threonine at its N-terminus, and
 CC a nucleophilic reagent able to cause cleavage of the intein to form a
 CC peptide bond between the extein C-terminus and synthetic peptide
 CC N-terminus through the formation of an activated ester or thio ester
 CC intermediate. (I) has antiinflammatory activity and prevents accumulation
 CC of pancreatic digestive enzymes, and prevents exocytic fusion of vesicles
 CC containing secretory enzymes of pancreas. (I) is useful for treating
 CC acute pancreatitis. The present sequence encodes the Clostridium
 CC botulinum BOTM/A neurotoxin prototoxin which is given in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 4835 BP; 1934 A; 517 C; 756 G; 1628 T; 0 other;

Query Match 54.6%; Score 727.6; DB 21; Length 4835;
 Best Local Similarity 72.2%; Pred. No. 2.4e-194;
 Matches 946; Conservative 0; Mismatches 364; Indels 0; Gaps 0;

OY	23	tgctacccttcacgaatataatcaagaacaatccaaatcccaatcccgtaaccggtcgt	82
Db	2945	tattctaattacttgtaataataataagaataataataataattcttatattgtaattgaat	3004
OY	83	acgaataccaatccacgtacgtacgacctgtctcgtctagctctccaaaatccaatcgggttcta	142
Db	3005	atggaagtaacttaatttaagacttattcgtatgtagtgcacaaaataaataattgttagta	3064
OY	143	aagttaacttcgatccgatctgcgaataagaatccagcttgtaactctgttaactctgtaattcca	202
Db	3065	aagtaaatatttgacccaatagataaaaatccaatccaatttaatttaatttaggaagtagta	3124
OY	203	aaatccgaagtatctccgaagaatcgtatcgtatacaaacctctgtacgaagaacattctcca	262
Db	3125	aaatggyggaatttttaaaaaatgcatatgtatataatagtatgtatgtatgtaaaaaatttagta	3184
OY	263	ccctcctcttgatccgtaatacccgaaaatactccaactccaatctctctgtgacaatagtataca	322
Db	3185	ctagctcttttgataagaattccttaagttattttaacagtaataagcttaataatagtata	3244
OY	323	ccatcatcaactgcatacgtgaaaacaattcctggttggaaagtatactctgtaactcggttgaa	382
Db	3245	caataataaattgtatagaaaataaattccagatcgtgaaagtatacctaattatattgtgtgaaa	3304
OY	383	tcacatcgtgacctgcgaagaactcgaagaataccaagaagtggtttatccaatattcttc	442
Db	3305	taactcgtgaccttcgaagaatacccggaataacaagaagtagtctttaataatcagtc	3364
OY	443	agatgatcaacatctcgtgaatacatcatcaatcgtcgtatcctgttgacatccaacaacatc	502
Db	3365	aaatgatataataataatcagatataataaaccgaatgggtttttgtaactatactaataata	3424
OY	503	gtctgtaataactccaanaatctatcatcaacgycgcgtctgtatcgacagaacccgatctcca	562
Db	3425	gattaaataactccaanaattatataataatggaagtataagatcaataaacaacatttcca	3484
OY	563	atcgcggtaacatccacagcgtcttaataaatoatcgttcaactatgtaacggttgtctgtgca	622
Db	3485	atttaggaataatcatcagctcagtaataataataatggttttaataatagtggtttgttagagata	3544
OY	623	ctccacgcctacatctgtgatcaataatctcatcatcgtctcgacacaagaactcgtgacaagaag	682
Db	3545	cacatagatatatttggataataaattcttaactcttttggatagaagatataatgaaagaag	3604
OY	683	aaatccaagaactcgtacgacacaccagttccaattctgtgatactctgaaagaactctcggggtg	742
Db	3605	aaatccaagaattctatagtaataatccaatccaatccagttattttaaagaacttttggggtg	3664
OY	743	actacccgcgtaagaacaaacgctactaaactcgtatgatctgtacagatccgaaacaaatcag	802
Db	3665	attcttatacatatagataaaacacatactaatatgtttaaatttataatgatcccaataaataatg	3724
OY	803	ttgacgttcaacaatgtagatcgcggtgtacatacgtacactctgaaaggtccgcggtgttctg	862
Db	3725	tcggtgttaaaataatgtagtgatttaggggtatataatgtatcttaaaaggccttagagtagcg	3784
OY	863	ttatgataccaacatactaccgtgaactcttcctcgtacacggttggttaaccaattcatcatca	922
Db	3785	taatgataccaacaacatttataatccaagtctgtataggggacaaaatttattataaa	3844
OY	923	agaaatacgcgctcgtgtaacaaagaacaaatatcgtctgcgaacatgatcgttgcatacatca	982
Db	3845	aaaaatactgcctcgtgaaataaagaataataatctgttgaataaataatgatcgttgtatata	3904
OY	983	atgctgtcagtttaagaacaagaatactacgcgtctgcgtccaaatgcttctcagcgtgtgtag	1042
Db	3905	atgtagtagttaaataataaagaataataagtttagctctataatgcataccacaggaagcgttag	3964
OY	1043	aaaagatcttgtctgtctcgtgaatacccggaactgttgataatactgtctcaggtagtgttaa	1102
Db	3965	aaaaatactcaatgagcatcttagaaatactcgtatgtaggaacttaagtaactgaatgtagtagtaa	4024
OY	1103	tgaaataccaagaacgacagggtatcatccaatacaaaaatgtaaatctgcgaagacaaca	1162

Db	4025	tgaagtcataaataatgatacaagatacaataaataagcaaatgatttacaagaataa	4084
Oy	1163	atggtaacgatcatcggttccatcggttccaccaggttcaacataatcgctaaactggtg	1222
Db	4085	atgggaatgatataagctttatagattatcattccatcggttataataatagactaaactag	4144
Oy	1223	cttcccaactggttcaatcgatcgatcgaaacgttccctccgcacgttgggttgccttg	1282
Db	4145	caagtaactgtgtataaagacaataagaaagactagtagagcttgggttgcctcaagg	4204
Oy	1283	agttcaccccggttgatgatcgatcggttggagacgttccgctgtaagaattc	1332
Db	4205	aattattcctgtagatgatgagtggtggggaagagccactgtaattatc	4254
RESULT	13		
AAZ87220			
AAZ87220	standard; DNA; 1317 BP.		
AAZ87220:			
08-MAY-2000	(first entry)		
DNA	encoding native BONT serotype A (BONTA) C-terminal fragment (Hc).		
Botulinum neurotoxin; heavy chain; BONT; serotype A;			
C-terminal fragment; Hc; Venezuelan equine encephalitis virus replicon;			
VEE; botulism; vaccine; diagnosis; drug screening; ds.			
Clostridium botulinum.			
Key	Location/Qualifiers		
CDS	1..1317		
/*tag=	a		
/product=	"Native botulinum neurotoxin serotype A		
(BONTA) heavy chain C-terminal fragment"			
/transl_except=	(pos:1294..1314, aa:gly)		
MO200002524-A2.			
20-JAN-2000.			
09-JUL-1999;	99MO-US15570.		
10-JUL-1998;	98US-0092416.		
12-MAY-1999;	99US-0133870.		
(USME-) US MEDICAL RES INST INFECTIOUS DISEASES.			
Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh WT, Smith L;			
WPI: 2000-160827/14.			
P-PSDB; AAY77142.			
Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum			
toxin serotypes A-G, is used for inducing an immune response against			
botulinum -			
Example 3; Page 52; 54pp; English.			

CC effects. The vaccine currently used against botulism is dangerous
 CC and expensive to produce, and contains formalin, which is very painful
 CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
 CC the 7 serotypes are represented in the formulation. The novel vaccine
 CC of overcomes these problems, as it is easily purified, and available in
 CC large quantities. It is also expressed in the lymph nodes for a better
 CC immune response. The present sequence represents DNA encoding native
 CC BONTA heavy chain C-terminal fragment (HC) used in an exemplification of
 CC the present invention.

XX
 CC
 SQ Sequence 1317 BP; 560 A; 108 C; 211 G; 438 T; 0 other;

Query Match 54.5%; Score 726.4; DB 21; Length 1317;
 Best Local Similarity 72.3%; Pred. No. 2.8e-194;
 Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 23 tgtctacctgaatacatcaagaacatcatcaatccctccatcctgaacgtcgct 82
 DB 14 tctctacattctgaatataatgaatattatcttctattcttgaatttaagt 73
 QY 83 agcaatccaatcactgtagcagcctgctcgtctgaagcttccaataatcaatcgtftcta 142
 DB 74 atgaaagtaactcatcttaataagactctatcagtagtgcacaataaataatattgtagta 133
 QY 143 aagtaactcgcagtcgacgacaagaatcagatccagctgcttccaatcctggaatctcca 202
 DB 134 aagtaaatcttgatccaatagataaataatcaatcaatattatattagaagtagta 193
 QY 203 aatcgaagtatacctgaagaatgctatcgtatcacactctatgtiagcaaaacttccca 262
 DB 194 aattgaggttaattttaaaaaaagctatgtatataatagtagtattgataaaatttagta 253
 QY 263 cctctctcgtgagtcgtatccggaataatttcaactccatccctcgaagaatgtatata 322
 DB 254 ctgagcttttgtagaagaatctcctaagatttcaacgataagtcctcaaatatgataata 313
 QY 333 ccatcatcaactgctgagaaacaatctggttggaaagatctctgaactacgtgtgaaa 382
 DB 314 caataataattgtatgagaaataatcagagtggaagtatcacttaatttgtgtgaaa 373
 QY 383 tcaatctggaactcgtcaggaacatcaggaatacaacagcgtgttgaatcaataactctc 442
 DB 374 taatctggaacttcaaggaatcactcaggaataaacaagaagtagtattttaatacagtc 433
 QY 443 agatgatcaacatcctcgtacatcaatcgcgtgagatctcgttaccatcacacaataac 502
 DB 434 aatgattcaatataatcagattatataaacaagatgattttgttaactatcacataataa 493
 QY 503 gtctgaataactccaataatctacatcaacgagcgtctgtatcgaccagaaacagatctcca 562
 DB 494 gattaaataacttaaaattatataatgaatgaagattatagatcaaaaacaaatttcaa 553
 QY 563 atctgggttaacatccagctcttaataacatcattcaactgtgaaggtgtcgtgtgaca 622
 DB 554 atttggaattatctatcagtagtaataataatgtttaattagatggtgtgtgagaca 613
 QY 623 ctccagcgtacatcctggaatcaacttcaatcgttgaagaagaagcagaagaagaag 682
 DB 614 caatagatataatttggataaataatttcaatcttcttgaataaggaattcaagaagaag 673
 QY 663 aatatcaagacctgtagcaacaacagctccaattctgtgtctcgtgaagaagctctggtgtg 742
 DB 674 aatatcaagattatataatgaataatcaatcaatcaggtatatttaaaagccttgggtgtg 733
 QY 743 actaccctgagtagcaacaacccgtatctacatgctgaatctgtacgactcggaacaataag 802
 DB 734 attattcaaatatgataaaccataactatgtttaattatattatgaccaataataatag 793
 QY 803 ttgaagtagcaaatgtagtaccggtttacatgtacttgaaaggtccggtgtgtctg 862
 DB 794 tctgagtagcaaatgtagtaccggtttacatgtacttgaaaggtccggtgtgtctg 853

QY 863 ttatgactacaacatctactcgaactcttccctgtaccgtgtgtaaccaatcatcatca 922
 DB 854 taatgctacaacacattatttaattcaagtgttgatagggggacaattatataaa 913
 QY 923 agaatacgcgtctgtgtacaagagacaaratcgttcgcacaatagatcgttatataca 982
 DB 914 aaaaataatcctcgtggaataaagaataatattgttgaataatgatcgttatataaa 973
 QY 983 atgtgtgtagtaagaacaagaataaccgctcgtcgtcgaactatcttccaggtgtgtag 1042
 DB 974 atgtagtagttaaaaataagaatagataggttagctactaatgtacacagcagcaggtag 1033
 QY 1043 aaaaagatctgtcgtcgtcgtcgaatacccggaagctgtgtgaatcgttcacagtagtga 1102
 DB 1034 aaaaataactaagtgattaggaataatcccgatgtgtaggaataatcgaagtagtaaa 1093
 QY 1103 tgaataccaagaagcagcaggtatcactcaacaatgtcaaatatgaatctgcaggaaca 1162
 DB 1094 tgaagtcaaaataatgtacaaggaatacaataataatgtcaaaatgtattcaagaatata 1153
 QY 1163 atgttagcagatattcgtttcattcgttcacacagttcaacatattcgttaactgtgtg 1222
 DB 1154 atgggaatgtatagagcttataagattcattcaatgaattatataatagctaaactagta 1213
 QY 1223 ctcccaactggtacaaatcgttcagatcgaacggttccctcgtcgaactcgtgtgtgtg 1282
 DB 1214 caagtaattggtatataatagacaataatagaagaatctagtagtactgtgtgtgtgtg 1273
 QY 1283 agttaccccggttgaatgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1326
 DB 1274 aatttattcctgttagatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1317

RESUME 14
 AAZ87218
 ID AAZ87218 standard; DNA; 2532 BP.
 XX
 AC AAZ87218;
 XX
 DT 08-MAY-2000 (first entry)
 XX
 DE DNA encoding native botulinum neurotoxin serotype A (BONTA).
 XX
 KW Botulinum neurotoxin; heavy chain; BONT; serotype A;
 KW Venezuelan equine encephalitis virus replicon;
 KW VEE; botulism; vaccine; diagnosis; drug screening; ds.
 XX
 OS Clostridium botulinum.
 XX
 FH Key
 FT CDS 1..2532
 FT Location/Qualifiers
 FT /tag= a
 FT /product= "Native botulinum neurotoxin serotype A
 FT (BONTA) heavy chain"
 FT /transl_except= (pos:2509..2529, aa:Gly)
 PN WO200002524-A2.
 PD 20-JAN-2000.
 XX
 PF 09-JUL-1999; 99WO-US15570.
 XX
 PR 10-JUL-1998; 98US-0092416.
 PR 12-MAY-1999; 99US-0133870.
 XX
 PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
 XX
 PI Lee JS, Pushko P, Smith JF, Parker M, Dertbaugh MT, Smith LJ;
 DR WPT: 2000-160827/14.
 DR P-PSDB: AA177140.
 XX
 PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 17:11:08 ; Search time 139.08 Seconds
(without alignments)
2352.487 Million cell updates/sec

Title: US-09-611-419A-1

Perfect score: 1332
Sequence: 1 gaatcgaacagatgcgtctt.....aacgtccctgtlaagaattc 1332

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCrTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	1311	98.4	1330	1	US-08-480-604A-22	Sequence 22, App	
2	1311	98.4	1330	2	US-08-405-496A-22	Sequence 22, App	
3	1311	98.4	1330	4	US-08-915-136-22	Sequence 22, App	
4	1311	98.4	1402	1	US-08-480-604A-25	Sequence 25, App	
5	1311	98.4	1402	2	US-08-405-496A-25	Sequence 25, App	
6	1311	98.4	1402	4	US-08-915-136-25	Sequence 25, App	
7	726.4	54.5	3891	1	US-08-480-604A-27	Sequence 27, App	
8	726.4	54.5	3891	2	US-08-405-496A-27	Sequence 27, App	
9	726.4	54.5	3891	4	US-08-915-136-27	Sequence 27, App	
10	166.4	12.5	1359	1	US-07-618-312A-3	Sequence 3, App	
11	166.4	12.5	1359	1	US-08-280-228-3	Sequence 3, App	
12	74.2	5.6	1359	1	US-07-618-312A-1	Sequence 1, App	
13	74.2	5.6	1359	1	US-08-110-786A-7	Sequence 7, App	
14	74.2	5.6	1359	1	US-08-280-228-1	Sequence 1, App	
15	74.2	5.6	1858	1	US-08-668-381A-6	Sequence 6, App	
16	60	4.5	5511	3	US-08-928-361B-2	Sequence 2, App	
17	60	4.5	5511	3	US-08-928-361B-1	Sequence 1, App	
18	57	4.3	5163	3	US-08-700-651-1	Sequence 1, App	
19	57	4.3	5163	3	US-08-928-361B-4	Sequence 4, App	
20	57	4.3	5318	3	US-08-700-651-2	Sequence 2, App	
21	57	4.3	5318	3	US-08-928-361B-3	Sequence 3, App	
22	40	3.0	1690	1	US-08-276-452A-24	Sequence 24, App	
23	40	3.0	1690	2	US-08-798-744-24	Sequence 24, App	
24	37.8	2.8	2277	1	US-08-676-967-2	Sequence 2, App	
25	37.8	2.8	2277	1	US-08-676-967-2	Sequence 2, App	
26	37.8	2.8	2277	2	US-09-098-487-2	Sequence 2, App	
27	37.2	2.8	18596	4	US-09-318-448-11	Sequence 11, App	

28	36.8	2.8	1430	1	US-08-276-452A-25	Sequence 25, Appl
29	36.8	2.8	1430	2	US-08-798-744-25	Sequence 25, Appl
30	34.6	2.6	7218	1	US-08-232-463-14	Sequence 14, Appl
31	34.6	2.6	2094	3	US-08-714-918-87	Sequence 87, Appl
32	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
33	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
34	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
35	34.4	2.6	951	1	US-07-855-412B-2	Sequence 2, Appl
36	34.4	2.6	951	2	US-08-308-887A-2	Sequence 2, Appl
37	34.4	2.6	951	3	US-08-881-094-2	Sequence 2, Appl
38	34.2	2.6	198	5	PCT-US95-10668-1	Sequence 1, Appl
39	34.2	2.6	198	5	PCT-US95-10668-2	Sequence 2, Appl
40	34.2	2.6	198	5	PCT-US95-10668-3	Sequence 3, Appl
41	34.2	2.6	198	5	PCT-US95-10668-4	Sequence 4, Appl
42	34.2	2.6	3292	1	US-07-814-964-12	Sequence 12, Appl
43	34.2	2.6	3292	1	US-08-258-442-12	Sequence 12, Appl
44	34.2	2.6	3292	1	US-08-328-809-7	Sequence 7, Appl
45	34.2	2.6	3292	5	PCT-US92-11107-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-480-604A-22
; Sequence 22, Application US/08480604A
; Patent No. 5736139
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHYE, NISHA V.
; APPLICANT: FIRDA, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,604A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422,711
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,496
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLTA, DIANE E.
; REGISTRATION NUMBER: 40,027

```

? REFERENCE/DOCKET NUMBER: OPHD-01763
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8610
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 22:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1330 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1314
US-08-480-604A-22

```

Query Match	98.4%	Score 1311;	DB 1;	Length 1330;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1311; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]

OY	796	aaatacgttcgcgcacaacatgtatgcgcggttaacatgtacatgtgaagctccgct	855
Db	787	AAATACGTTGACGTCACACATGTAAGTATCCGGGTTACATGTAAGTGAAGTCCGGCT	846
OY	856	ggcttcgttatgtactaccaacatctactctgaactcttcctctgaacgtygtaccaaatc	915
Db	847	GGTCTCTGTAATGACGTACCAACATCTTACCTGATGACTTCCCTGTACCGTGTACCAATTC	906
OY	916	atcatcaagaatctcgctctctgtgaacaagaacataatcgttcgcaacaatgatctgtga	975
Db	907	ATCATCAAGAAATACGGCGTCTGGTATACAGACAAATATCTTTCGACAAATATCTGTGA	966
OY	976	tacataatgtcttgatgttaagaacaaagaaataacgctctggtctacaaatgcttcgaagct	1035
Db	967	TACATCAATGTTGTAAGTAAGAACAAAGATATACCGTCTGGCTACCAATCTTCTCAGCGCT	1028
OY	1036	ggtgtagaagaagatctgtctgtcctctgtgaatcccggaagttgtgtatctgtctcaagta	1099
Db	1027	GGTGTAGAAAGATCTTGCTGCTCTGTGAATATCCGGACGTTGTATATCTGTCTCAGGTA	1088
OY	1096	gtctgaatgaatccaagaacagcaagagatcatcaatacaaatgtcaaaatgtaatctgcag	1155
Db	1087	GTTGTAATGAATCCAGAACGACCAAGATGATCATATACAAATATGCAATATGATCTGCAG	1146
OY	1156	gacacaaatggttaacgataatccgttttcatctggtttccacaggttaacaaatatcgctaa	1211
Db	1147	GACCAACATGGTAAACGATTCGGTTTCATGGTTTCCACAGATTCACAAATATCTGCTAAA	1200
OY	1216	ctggttgccttccaaactgtatcaatcgcgcagatcgaaacgcttcctctgcacactggtgtgc	1275
Db	1207	CTGGTGTCTTCCACTGATCAATCGTCAATGGAAGCTTCTCTCTGCACTCTGGGTTGCT	1266
OY	1276	tcttggaagatcaatcccggttgatgatcaggtttgggggtgaacgctgcgtctga	1326
Db	1267	TCTTGGAAGATTCATCCCGGTTGATGACGCTTGGGGGTGAACGCTGCCCTGTAA	1317
RESULT 2			
US-08-405-496A-22			
Sequence 22, Application US/08405496A			
Patent No. 5919665			
GENERAL INFORMATION:			
APPLICANT: WILLIAMS, JAMES A.			
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM			
TITLE OF INVENTION: NEUROTOXIN			
NUMBER OF SEQUENCES: 30			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: MEDLEN & CARROLL, LLP			
STREET: 220 MONTGOMERY STREET, SUITE 2200			
CITY: SAN FRANCISCO			
STATE: CALIFORNIA			
COUNTRY: USA			
ZIP: 94104			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOOS			
SOFTWARE: Patentln Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/405,496A			
FILING DATE: 16-MAR-1995			
CLASSIFICATION: 424			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 08/329,154			
FILING DATE: 25-OCT-1994			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 08/161,907			
FILING DATE: 02-DEC-1993			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 07/985,321			
FILING DATE: 04-DEC-1992			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 07/429,791			

```

; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-01308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1314
; US-08-405-496A-22

```

```

Query Match          98.4%; Score 1311; DB 2; Length 1330;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 16 cgtcgcctgtctacccctcctgaatacatcaagaacatcatcaatccctccatccgaac 75
DB 7 CGTCTGCTGTCTACTCTTCTCATATCATCAAGAACATATCATCTCTCTGAAAC 66
OY 76 ctgcgcctgaatcccaatccatctgcacgtctcgtctacgtctccaaatccaatc 135
DB 67 CGCGCTGTGAGATCCCAATCACCATCGATCGACGTCTCGCTACGCTTCCAAATTCACATC 126
OY 136 ggtttaaagttaacttcgatccgatccgaagaatcaatccagctgttcaatctcgaa 195
DB 127 GGTCTAAAGTTAACTTCGATCCGATCGACGAAGAATCAAGATCAGCTGTTCAATCTGAA 186
OY 196 tcttccaaatcgaaagttactcctgaagaatgtctatcgtatcaaatctgtacgaaac 255
DB 187 TCTTCCAAATGGAAGTATCTCTGAAGATGCTATGCTACACCTATGTAAGAAAC 246
OY 256 tcttccacctctctgtgacccgatccgaataacttcaactccatctctgaacaat 315
DB 247 TTCTCCACCTCTCTGTGATCCGATCCGAAATCTTCAACTCCATCTCTGTGAACAT 306
OY 316 gaatacaccatatacaactgcatagtgaagaacatctcgtgtgaagaatctctgaactc 375
DB 307 GAATACACCATCATCAACTGATGAAACAAATCTGTGGAAGATATCTTGAACCTAC 366
OY 376 ggtgaatcatctggaactctgcaggaacactcaggaatacaacagcgtgtgtatccaa 435
DB 367 GGTGAATATATCTGACTCTGACAGACACTCAAGAAATCAACAGCGTGTGTATTCAAA 426
OY 436 tactctcagaatcaacaactctgtacatacatcaatcgtgtgatctctgaaccatcc 495
DB 427 TACTCTCAGATGATCAACATCTCTGACTACATCAATCGCTGATCTTCTGTACCATCAC 486
OY 496 aacaatcgtctgaataaccccaataatctacatacaagcgcgtctgtatgaacgaagac 555
DB 487 AACCAATCGTGAATTAACCCAAAATCTATCAATCAAGCGCTGTGATCAACGAAGACG 546
OY 556 atctccaatctgtgtaaacatccagcgtcttaataacatcatcttcaaatctgaacgtgt 615
DB 547 ATCTCCAATCTGGTAACTCAACGCTTCTTAATTAATCATGTTCAAACTGACGCTTGT 606
OY 616 cgtgaactaccgcgtacatctgtgatcaaatcttcaatctgttcgaagaagaactgaac 675
DB 607 CGTGACACTCACCGCTACATCTGAGATCAATCTTCAATCTGTTCGACCAAGAACTGAC 666
OY 676 gaagaagaatacaagaactgtacagcaacagttcaatcttgatctctgaagaagcttc 735
DB 667 GAAAAAGAAATCAAGACCTGTACGACACCAAGTCCAAATCTGTGATCTCTGAAAGACTTC 726

```

```

OY 736 tgggtgactactcgtcagtcagcaaacccgtactacatgtgaatctgtacgatccgaac 795
DB 727 TGGGGTACTACTTGCAGTACGACAAACCGTACTACATCTGAAATCTGTACGATCCGAAC 786
OY 796 aaatcgttgaacgtlcaacaatgttagtalcgcggtgttacatgttaccgtgaagctcc 855
DB 787 AAATACGTTGAGCTCAACATGTAGGTATCCGGGTACATGACTGAAAGTCCGCGCT 846
OY 856 gttctgtatgactaccacaatctacatcgaactctctcccttacggtgttccaaatcc 915
DB 847 GGTCTGTTATGACTACCAACATCTACCTGAACCTTCTCTGTACCGGTGACCAAAATTC 906
OY 916 atcatcaagaataacgcgtctgttgaagaagaacatctgttcgaacaataatcgtgta 975
DB 907 ATCATCAAGAAATACCGCTGTGTATCAAGGACAAATATGTTCCGCAACATGATGCTGTA 966
OY 976 taatcaatgttctgttgaagaagaataacgcgtctgtgtcgtacaaatgtcttcagct 1035
DB 967 TACATCAATGTTGTACTTAAGAAACAAAGATACCGTCTGCTACCAATGCTTCTCAGGCT 1026
OY 1036 ggtgtgaagaagaatctgtctgtctgtgaaatcccggaacgttgtatctgtctcagta 1095
DB 1027 GGTGTGAAAAGATCTGTCTGTCTGTGAATTCGCGACGTGTGTATCTGTCTCAGTA 1086
OY 1096 gttgtaataatcccaagaacgacaggtatcactcaacaatgcaaatgaatctgcag 1155
DB 1087 GTTGTAAATGAATCCAGAAACGACGAGGTATCATTACAAATGCAAAATGAATTCGACG 1146
OY 1156 gacaacaatgtaagaatcgtgttcaatcgtgttccaccagttcaacaatatcgtctaa 1215
DB 1147 GACAAATAGTATGATGATTCGCTTCATCGGTTCCACAGTTTCAACATATTCGCTAAA 1206
OY 1216 ctgtgtgtcttccactggtgaacatcgttcgaatcgaactctcctcgtcactcgtgttc 1275
DB 1207 CTGTTGCTTCCACTGTGATCAATCTGATGAAATCCGATCGTCTCTGCTGCTGGGTTC 1266
OY 1276 tcttgggaatcaccctcgttgaacggttgggtgaacggttcggtcgtcgtga 1326
DB 1267 TCTTGGAGTATATCCCGGTGTGATGACGGTGGGGTGAAGTCGCTGTATA 1317

```

```

RESULT 3
US-08-915-136-22
; Sequence 22, Application US/08915136
; Patent No. 6290960
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHYE, NISHA V.
; APPLICANT: FIRCA, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; PREVENTION OF C. DIFFICILE DISEASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,136
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/480,604
; FILING DATE:

```

Query Match	Score	DB	Length
Best Local Similarity	100.0%	Pred. NO. 0;	
Matches 1311;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;

Query	Match	Score	DB	Length
QY	16	98.4%	Score 1311; DB 4;	Length 1330;
Db	7			
QY	76			
Db	67			
QY	136			
Db	127			
QY	196			
Db	187			
QY	256			
Db	247			
QY	316			
Db	307			
QY	376			
Db	367			
QY	436			
Db	427			
QY	496			

Db	487	AACATCGTGTGAATTAATCCAAAATCTACATCAACGGCGTCTGATGCAGCAACGAACCG	546
Qy	556	atcccaactctgggttaacatccacgctctcaataacatcatgttcaactggaagtggt	615
Db	547	ATTCACCAATCGGGTAACTACACACGGCTTCTAAATACATCATGTCTAAACTGGACGGTGT	606
Qy	616	cgtaacctaacccgtctacatcgtatgataacaaattactcaatctgtttgnaaagaactgaac	675
Db	607	CGTGACACTCACCCCTACATCTGGATCAATTAATTCAATCTGTTGCACAAAGAACTGAAC	666
Qy	676	gaaaagaataacaaagacccgttaccgaacacagttccaaattcgtlaacctgaagaacttc	735
Db	667	GAAAAGAAATCAAAAGACCTGTACAGAACACAGCTCCAAATTCGTGTCTGAAAGACTTC	726
Qy	736	tggggtgactactcctgcagttacgacaaacggtactacatgctgaatctgtacatccgaac	795
Db	727	TGGGGTGACTACCTCGACGTACGACAAACCCGTACTACATGCTGAAATCTGACATCGAAC	786
Qy	796	aaatacgttgaacgtcaacaaattgtagtataccgggttaacatgtaactggaaggtccgggt	855
Db	787	AAATACGTTGACGTCACCAATGTAGGTATCCGGGTACATGTACCTGGAAGGTCGGCT	846
Qy	856	ggctctgttgaatcacaacatcatctaacatctctccctgtacacggttgtaaccaatlc	915
Db	847	GGTCTCTGTATGACTTACCAACATCTACCTGACTCTTCCTGTACCGTGATGACCAATTTC	906
Qy	916	atcatalcaagaataacgcgtctctgtgtaacaaagacaatatcgttcgcaacaatgatcgtgla	975
Db	907	ATCATCAAGAAATACGGCGTCTGGTAACACAGCAATATCTGTCGCAACAAATGATCGTGA	966
Qy	976	tacatcaatgttctgaagttaagaacaaagaataccggtctggtctaccaatgtcttcaagct	1033
Db	967	TACATCAATGTTGATGTTAAGTAAGAACAAAGAAATACCGTGGCTACCAATCTTCTCAGGCT	1026
Qy	1036	ggcttagaagaagactctgtctgcctctggaaaatcccggaacgttbgtaatctgtcacaqta	1095
Db	1027	GGTGTAGAAAAGATCTTGTCTGCTCTGTGGAATCCCGAGCTTGTAATCTGTCTCAGTA	1088
Qy	1096	gtctlaatgaaatccaaagacgcacaggtgatcactaacaatgycaaaatgaaatcgtcacg	1155
Db	1087	GTTGTAATGAATCCAGAACGACACAGGGTATCATAACAAATGCAAAATGAAATCGACG	1146
Qy	1156	gacacaaatggttaacgatatcggtttcatcgtgtttccacagttcaacaaatatcgcttaa	1211
Db	1147	GACACAAATGTTAATCGATATCGGTTTCATCGTTTCCACACAGTTCAAATATCCCTAAA	1200
Qy	1216	ctggttctctccaactcgttacaatcgtcagaatcgaaacgttccctcgtcagctctgggttgc	1275
Db	1207	CTGGTGTGTTCCAACTGGTACAAATCGTCAATGGAACGTTCTCTCTGCACCTCTGGGTTGC	1266
Qy	1276	tcttggagatcaccccggttgaatgaacggttgggtggaacggtccgctgtlaa	1336
Db	1267	TCTTGGGAGTTCATCCCGGTGTGATGACAGGTTGGGGTGAAGCTCCGCTGTAA	1317
RESULT 4			
US-08-480-604A-25			
Sequence 25, Application US/08480604A			
Patent No. 5736139			
GENERAL INFORMATION:			
APPLICANT: KINK, JOHN A.			
APPLICANT: THALEY, BRUCE S.			
APPLICANT: PADHYE, NISHA V.			
APPLICANT: FIRCA, JOSEPH R.			
APPLICANT: STAFORD, DOUGLAS C.			
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND			
PREVENTION OF C. DIFFICILE DISEASE			
NUMBER OF SEQUENCES: 32			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: MEDLEN & CARROLL, LLP			
STREET: 220 MONTGOMERY STREET, SUITE 2200			
CITY: SAN FRANCISCO			
STATE: CALIFORNIA			


```

? GENERAL INFORMATION:
? APPLICANT: WILLIAMS, JAMES A.
? TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
? TITLE OF INVENTION: NEUROTOXIN
? NUMBER OF SEQUENCES: 30
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: MEDLEN & CARROLL, LLP
? STREET: 220 MONTGOMERY STREET, SUITE 2200
? CITY: SAN FRANCISCO
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 94104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/405,496A
? FILING DATE: 16-MAR-1995
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/329,154
? FILING DATE: 25-OCT-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/161,907
? FILING DATE: 02-DEC-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/985,321
? FILING DATE: 04-DEC-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/429,791
? FILING DATE: 31-OCT-1989
? ATTORNEY/AGENT INFORMATION:
? NAME: INGOLIA, DIANE E.
? REGISTRATION NUMBER: 40,027
? REFERENCE/DOCKET NUMBER: OPHD-01308
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8410
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1402 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1386
? US-08-405-496A-25

```

Query Match 98.4%; Score 1311; DB 2; Length 1402;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

? 16 cgtctgctgtctactcctcactgaatacatcaagaacatcatcaatcctcctgaac 75
? 79 CGTCTGCTGTCTACTCTCAGTAATCATCAAGAATCATCATCACTCATCTGAGAC 138
? 76 ctgctgaagaatccaatcaactcgtacgtctcgtctgaagcttccaataacac 135
? 139 CTGGGCTACGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATC 198
? 136 gttctcaaatgaatcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 195
? 199 GGTCTTAAGTTAATCTCGATCCGATCGACAAATCATCATCATCATCATCATCAT 258
? 196 tcttccaaatcgaagtatactcgaagaatgctatcgtatatacaactctatgta 255
? 259 TCTTCCAAATCGAAGTATCTCTGAAGAATCTATCTATCACTATATGTACGAAAC 318

```

```

? 256 tcttccacctctctctgtagtcgtagtaccggaataactcaactccatctctga 315
? 319 TTCTCCACCTCTCTCTGATCCGATATCCGAATTAATTAATTAATTAATTAATTA 378
? 316 gaataacacatcactcactgtagtgaataacaaatcttggttggaagaatctta 375
? 379 GAAATACACCATCATCACTGATGATGAATAATCTGTTGGTGAAGATATCTTAAC 438
? 376 ggtgaatcactcgtactcgaactcgaactcgaactcgaactcgaactcgaactc 435
? 439 GGTGAATCATCTGACTCTGACTCTGACTCTGACTCTGACTCTGACTCTGACTCT 498
? 436 tacttcaagatgaataactcctcgaactcgaactcgaactcgaactcgaactc 495
? 499 TACTCTCAGATGATCAATCTCTGACTCTGACTCTGACTCTGACTCTGACTCTG 558
? 496 aacaatgctcgaataactcgaataactcgaataactcgaataactcgaataact 555
? 559 AACATGCTCTGAATTAATCTCAAAATCTACATCAACGCGCTGATCGAACAAAC 618
? 556 atctccaatctggttaacatccacgctctcctataacatcgttcaactggaag 615
? 619 ATCTCCATCTGGGTATCATCTCAGCTCTCTAATTAATCATCTCAAACTGAGCT 678
? 616 cgtgaactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcga 675
? 679 CGTACACACTCACCCCTCATCTGATCATCAATTAATTAATTAATTAATTAATTA 738
? 676 gaaaagaatcgaagaactctcgaagaactcgaagaactcgaagaactcgaaga 735
? 739 GAAATGAATCAAGACCTGATGACACCAAGTCCAAATTTGGTATCTTAAGACTTC 798
? 736 tgggtgactacacgctcgaactcgaactcgaactcgaactcgaactcgaactc 795
? 799 TGGGTGACTACCTGACGATGACAAACCGTACTGATGATGATGATGATGATGAT 858
? 796 aaatacgttgaactcgaactcgaactcgaactcgaactcgaactcgaactcga 855
? 859 AAATACGTTGACGTCACCAATGATGATGATGATGATGATGATGATGATGATG 918
? 856 ggtctgttgaactcgaactcgaactcgaactcgaactcgaactcgaactcga 915
? 919 GGTCTGTATGACTACCAATCTACTGATGATGATGATGATGATGATGATGATG 978
? 916 atcatcaagaatacgaactcgaactcgaactcgaactcgaactcgaactcga 975
? 979 ATCATCAAGAAATACGCTGATGATGATGATGATGATGATGATGATGATGATG 1038
? 976 taactcaatgctgtagtgaataacaaagaataacgctcgtctacaaatgctc 1035
? 1039 TACATCAATGTTGATGATTAAGAAACAAATACGCTGCTGCTACCAATGCTTC 1098
? 1036 ggtgtagaagaatctctcgtcctcgaactcgaactcgaactcgaactcgaact 1095
? 1099 GGTGTAGAAAAGATCTTCTGCTCTGGAATTCCTCGAGACGTTGGTATCTCT 1158
? 1096 gctgtaatgaatacgaactcgaactcgaactcgaactcgaactcgaactcga 1155
? 1159 GTTGAATGAATCAAGAAACGACGATGATGATGATGATGATGATGATGATG 1218
? 1156 gacaacaatgtagatgtagatgtagatgtagatgtagatgtagatgtagatg 1215
? 1219 GACACAAATGTTAAGATATGATGATGATGATGATGATGATGATGATGATG 1278
? 1216 ctggtctcgaactcgaactcgaactcgaactcgaactcgaactcgaactcga 1275
? 1279 CTGGTGTCTTCAACTGTTAATCTGATGATGATGATGATGATGATGATGATG 1338
? 1276 tcttggagatcattcccggttgaactcgaactcgaactcgaactcgaactcga 1326
? 1339 TCTTGGAGATTCATCCGCTGATGATGATGATGATGATGATGATGATGATG 1389

```



```

0Y 1156 GACAAcAtggtlaacgaAtcgggtttccatcggtttccaccagttcaacataAtcgctaa 1215
1219 GACAAcAATGtTAACGAtATATCGGTTTCATTCGGTTTCACACAGTTCAACAATATCCCTAA 1278
Db 1219 CTGGTTGGTTCCAACTCGTACAAATCGTCAGATCGAAGCAGTTCCCTCTGCGACACTCGGGTTGC 1338
0Y 1216 ctggttgcttccaaAtcggttaacAtcgtcagatcgaAttcgAttcctctgcacAtcgtgttgc 1275
1279 CTGGTTGGTTCCAACTCGTACAAATCGTCAGATCGAAGCAGTTCCCTCTGCGACACTCGGGTTGC 1338
Db 1276 tcttgaggatcAtccacgggttgcagtggttgggtggaacgtccgctgtaa 1326
1339 TCTTGAGGAtTCATCCCGTTTGATGACGCTGGGGTGAAcGTCcCTGTAA 1389
RESULT 7
US-08-480-604A-27
Sequence 27, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHAYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLTA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3891 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

[illegible]


```

Db 3428 TAAAGACTACAAACATTATTTAAATTCACGTTGTATAGCGGACAAATTTATTTATA 3487
Qy 923 agaaatcgcgttcgttaacaagaacatcgttcgaacaatgacgttatactca 982
Db 3488 AAAAATATGCTTCGGAATAAAGATATATGTTAGAAAATATGATCGTATATATTA 3547
Qy 983 atgtttagttaaagaacaagaatacgcgttcgttaccatgcttccaaagctgttag 1042
Db 3548 ATGTAGTAGTTTAAATAAAGAAATATAGTTAGCTACTAATGATCAGACGACGCGTAG 3607
Qy 1043 aaagatctgtctgcctcgtgaatccgcgaactgtgttaactctcagtaattgaa 1102
Db 3608 AAAAATACTAGTGTGATTGAATAATACCTGATGTAGAAATCTAGCTCAAGTAGTAGTAA 3667
Qy 1103 tgaatcacaagaacagcaggttaccatacaaatgcaaatgcaatcgtcgaaggaca 1162
Db 3668 TGAAGTCAAAAATGATCAAGATTAACAATAATGCAAAAATGATTTACAAGATATA 3727
Qy 1163 atgttaacagatcgttccatcgttcacccagttccaaacataatcgttaactgttg 1222
Db 3728 ATGGGAATGATATAGCTTTATAGATTTTCATCAGTTTAAATATATAGCTTAAGTAGTAG 3787
Qy 1223 ctcccaactggtacaaatcgttcagatcgaacgttctctcgcacgtcgtgtgtcttggg 1282
Db 3788 CAACTTAATTTGGTATTAATAGCAAAATAGAAAGATCTAGTAGGACTTTGGGTTGCTCATGGG 3847
Qy 1283 agtcaaccgcgttcgtgaacggttgggtgaacgtccgctgtaa 1326
Db 3848 AATTATTTCCGTGATGATGATGATGGGAGAAAGCCACCTGTAA 3891

```

RESULT 9
US-08-915-136-27
Sequence 27, Application US/08915136
Patent No. 6290960

GENERAL INFORMATION:

```

APPLICANT: KINK, JOHN A.
APPLICANT: THALEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk.
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,136
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,604
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3891 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3888
US-08-915-136-27

```

Query Match 54.5%; Score 726.4; DB 4; Length 3891;
Best Local Similarity 72.3%; Pred. No. 7.9e-204;
Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

```

Qy 23 tgtctacccatcgtgaatcacatcaagaacatcattacatccatccatcgtacgtcgt 82
Db 2388 TATCTACATTTACTGGAATATATATTAAGATATATTAATTAATTAATTAATTAATTAAT 2647
Qy 83 acgaatcaatcgcgtcgtacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 142
Db 2648 ATGAAGAATGATTTATTAATAGACTTATCTAGTATGATGATCAAAATATTAATTTGTA 2707
Qy 143 aagtaacttcgatcgcgtcgaagaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 202
Db 2708 AAGTAAATTTTGATCCAAATAGATTAATAATTAATTAATTAATTAATTAATTAATTAAT 2767
Qy 203 aaatcgaatcgtcgtgaagaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 262
Db 2768 AAATGAGGATTAATTTTAAATAATCTATTTGATATATATATATATATATATATATAT 2827
Qy 263 cctcccttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 322
Db 2828 CTACCTTTTGGATTAAGATTTCTTAAGTATTTTAAACAGTATAGCTTAATTAATTAATTA 2887
Qy 323 ccatacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 382
Db 2888 CAATAAATTAATTTGATGGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2947
Qy 383 tcatctgacctcgtcgaagacatcgaagacatcgaagacatcgaagacatcgaagacatcgtc 442
Db 2948 TAACTCGACTTTTACAGGATCTCAGGAAATTAACAAAGAGTGTGTTTAAATTAACAGTC 3007
Qy 443 agatgatcaaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 502
Db 3008 AAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3067
Qy 503 gtcgtgaatactccaatactacatacgaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 562
Db 3068 GATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3127
Qy 563 atctgggttaacatccagctcttaataacatcgaatcgaatcgaatcgaatcgaatcgaatc 622
Db 3128 ATTTAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3187
Qy 623 ctcaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 682
Db 3188 CACATAGATTAATTTGATTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 3247

```

Qy	663	aaatccaagaccgtgtaacgaaccagttcccaatctcgtatcctgtaagaactctcgtgggtg	742
Db	3248	AAAATCAAGATTTTATATAGTAATCAATCAAAATTCAGATATTTTAAAGACTTTTGGGGTG	3307
Qy	743	actaccgcagtcgaacaaaccgtaactatctgaatctgtaatctgtaactccgaacaaatcg	802
Db	3308	ATTATTTCATATATGATTAACCCCTACTATATGTGTTAAATTTATATGATCCAAATATATATYG	3367
Qy	803	ttgacgcacaacatgtaagatctccgcggttacaatgtaactggaagaagtcgcggtgtctg	862
Db	3368	TGCGATGAATTAATATGATAGATATATGAGGTTATATGATCTTAAAGGGCTGAGAGGTACG	3427
Qy	863	ttatgaatcccaacatctcaactcgtgaactcttccctgtaccggtgtaaccaatcatca	922
Db	3428	TAAAGACTCAAAACATTTATTTTAAATTCAGATGTTGTATGAGGGGGCAAAATTTTATATA	3487
Qy	923	agaatatcgctctggttaaccaagaacatatacgcttcgcaacaatgatcgtgtatacatca	982
Db	3488	AAAATATTCGCTTCTCGAAATTAAGATATATGTGTTATGGAATATATATGCTGATATATTA	3547
Qy	983	atgttgtagttaagaacaaagaatatcgcgtctcgaatcgaatgcttctcaagcttggttag	1042
Db	3548	ATGTAGTAGTTAAAAATTAAGATATATAGGTTAGCTACTATGCATCACAGCAGCGAGTAG	3607
Qy	1043	aaaagactctgtctgtctcgtggaatcccggaagcttggtaatctgtcctcaagtgatgtgaa	1102
Db	3608	AAAAAATACTAAGTCGCTTATGAATATCCGTATGTAGGAATTTAAGTCAAGTAGTAGTAA	3667
Qy	1103	tgaatccaagaacagcaccgaaggtatacctaacaacatgcaaatgtaatactgcaggaacaa	1162
Db	3668	TGAAGTCAAAAATATGATCAAGGAATTAACAATTAATATGCAAAATGAAATTTTACAAGTAAFA	3727
Qy	1163	atggtaacgatcgcgtttcatcgtgttccaccggttcaacaatatcgctaactgtygtg	1222
Db	3728	ATGGGAATGATTAAGCGCTTATATGAGATTTCAATCAGTTTAATTAATATACCTAAACTAGTAG	3787
Qy	1223	cttccaactgtgtaacatcgtaagatcgaaatcgcgttccctctgcactcttggtgtgctctgg	1282
Db	3788	CAACTAATTTGTATTAATAGCAAAATAGAAAGATCTAGTAGAGCTTTGGGCTGCTCATGGG	3847
Qy	1283	agttcaccccggttgatgaagcgttgggtggaacgctccgctgtaa	1326
Db	3848	AATTTATTCCTGATAGATGATGATGGGAGAAAGCGCACTGTAA	3891
RESULT 10			
US-07-618-312A-3			
Sequence 3, Application US/07618312A			
Patent No. 5389540			
GENERAL INFORMATION:			
APPLICANT: Makoff Dr, Andrew J			
APPLICANT: Romanos Dr, Michael A			
APPLICANT: Clare Dr, Jeffrey J			
APPLICANT: Fairweather Dr, Nell F			
TITLE OF INVENTION: VACCINES			
NUMBER OF SEQUENCES: 13			
CORRESPONDENCE ADDRESS:			
ADDRESS: 14th Floor			
STREET: 2200 Clarendon Boulevard,			
CITY: Arlington,			
STATE: Virginia			
COUNTRY: U.S.A.			
ZIP: 22201			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentin Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/07/618,312A			
FILING DATE: 19910516			
CLASSIFICATION: 424			

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  GB 8926832.0
3      FILING DATE:  28-NOV-1989
4
5      PRIOR APPLICATION DATA:
6      APPLICATION NUMBER:  GB 9006097.1
7      FILING DATE:  17-MAR-1990
8      ATTORNEY/AGENT INFORMATION:
9      NAME:  Crawford Mr, Arthur R
10     REGISTRATION NUMBER:  25,327
11     REFERENCE/DOCKET NUMBER:  510-51
12     TELECOMMUNICATION INFORMATION:
13     TELEPHONE:  0101 703 8750400
14     TELEFAX:  0101 703 5253468
15     TELEX:  200797 NIXN UR
16
17     INFORMATION FOR SEQ ID NO:  3:
18     SEQUENCE CHARACTERISTICS:
19     LENGTH:  1359 base pairs
20     TYPE:  NUCLEIC ACID
21     STRANDEDNESS:  double
22     TOPOLOGY:  linear
23
24     MOLECULE TYPE:  cDNA
25     HYPOTHETICAL:  NO
26     ANTI-SENSE:  NO
27     ORIGINAL SOURCE:
28     ORGANISM:  Clostridium tetani
29     FEATURE:
30
31     NAME/KEY:  CDS
32     LOCATION:  1..1356
33
34     US-07-618-312A-3

```

Query Match	Similarity	Score	16.4	DB 1:	Length	1359;
Best Local	Similarity	52.3%;	Pred.	No.	2.9e-39;	
Matches	486;	Conservative	0;	Matches	411;	Indels 33; Gaps
49	aacatcatcaaacctccatccctgaacaccgcgcgcgtacgaatccaatcaactgaatgcacgtc	108				
49	ATTCGTGAATAAAAGTATACCACTTCTGAACCTTGAGCATCAACAGATATATTATTCGCCGATC	108				
109	tctgcgtacgcttccaaaatacaacatcggttttaagttaacttgcgatcgg---atcgac	165				
109	TCTGGTTTCACACTCCCTCTGTTATCAGATATCCAGATGCTCAATTGGTGGCGGACATCAAC	168				
166	aagaatcagatccagcgtgttcaactcgtgaactcttccaactccaagaatgaatgtaactcgtgaagaat	225				
169	GGCAAAAGCATATCCACTGGTTAAACAAGAAATTTCTGAAGTATATCGACACAAGGCCATG	228				
226	gctatcgtatatacaactatgtatgcgaaaactcttccaactcttctgtgatccgtatcccg	285				
229	GACATCGAATATCAAGACATATGTTCAACAACATCACCGTTAGGTTCTGGCTGGCGGCTTCCG	288				
286	aaatacttcaactcatc---tctcgaacaatgaataccaatcaactcgtc	336				
289	AAAGTTTCTGCTTCCCACTCGAAGACAGTATGGGCATTAAGATTAATCTCCATCATCACTCT	348				
337	atggaanaaat-----tctggttggaagaatctcttgaactcgtatga	381				
349	ATGAGAAACATCTCCCTGTCATTCGGCTCTGGTTGGTGTGTTCCCTGTAAGGATTAACAC	408				
382	atcatcttgacctcgaagacactcgaagaaatcaacaacagcgtgtgtatcaataatct	441				
409	CTGATCTGAGACTCTAANAAGACTCCGGCGGCGGAAGTTTCGTCATATCAATCACTTCCGCGACCTG	468				
442	cagatgatcaaatctcttgactatacaatcgcgttgatcttcgttaccataccaacaat	501				
469	CCGGCAAAAGTTATCAAGCGTACCTGGCTTAACAATGGGTTTTCATCATCACTTACTACATGAT	528				
502	cgctcgaataactccaataatctacatcaacagcgctgtgatctgaccagaacacgattcc	561				
529	CGTCTGCTTCTGCTAACCTGTACATCAACAGCGCTTGTGATGGCTCCGCTGAATATCACT	588				
562	aatctggttaaatcaacgctctctaataacaatcatgttcaaaccttgagacggttgcgtgac	621				

Db 589 GGTCTGGGCGCTATCCGTGAGACAAACATCATCTTTAAGCTGGACCGTTCACAC 648
QY 622 accaacgcgtacatcttgatcaatacttcaatctgttcgacaaagacgtgaacga 681
Db 649 AACACACAGTACGTATCCGACACAGTATCCGATCTTTGCAAGACACTGAAACCG 708
QY 682 gaaatcaagaactgtacgaacacagltccaaatctgtatcttcgaaagacttcg 741
Db 709 GAGATCGAAAAAAGTGTATACACTACTGCTATACACTTCTGCTGCTGCTGCTG 768
QY 742 gactactgtgacgaacacacgttactatactgtcaatctgtacgaatccgaacaa 801
Db 769 AACCCGCTGGCTTACGACAGGAATATATCCGGTACCTTCTAAGCTTAAAGAC 828
QY 802 gttgacgtcaacaatgtaggtatccgcggtgacgtgacgtgacgtgacgtgacgt 861
Db 829 GTTCAGCTGAAAA-----CATCAGTACTATACGTACTGACACAGGCGCTCTAC 882
QY 862 gttatgactcaacaatctacgtgaactcttccctgtacgtgtgtacacaaatcaat 921
Db 883 ACTAACGCGTAAACTGAACTACTACTACCGACGCTGTACAAAGCGCTGAAATTCATC 942
QY 922 aagaatacgcgtctgttaacaaagacaat 951
Db 943 AAACGCTACACTCCGAACACGAATCGAT 972

RESULT 11

US-08-280-228-3

Sequence 3, Application US/08280228

Patent No. 5571694

GENERAL INFORMATION:

APPLICANT: Makoff Dr, Andrew J

APPLICANT: Romanos Dr, Michael A

APPLICANT: Clare Dr, Jeffrey J

APPLICANT: Fairweather Dr, Neil F

TITLE OF INVENTION: VACCINES

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHUYE P.C.

STREET: 1100 No. 5571694th Glebe Road

CITY: Arlington,

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,228

FILING DATE: 25-JUL-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/618,312

FILING DATE: 27-NOV-1990

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 8926832.0

FILING DATE: 28-NOV-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9006097.1

FILING DATE: 17-MAR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Wilson, Mary J.

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 117-163

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

TELEX: 200797 NIXN UR

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 1359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Clostridium tetani
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1356
US-08-280-228-3

Query Match 12.5%; Score 166.4; DB 1; Length 1359;
Best Local Similarity 52.3%; Pred. No. 2,9e-39;
Matches 486; Conservative 0; Mismatches 411; Indels 33; Gaps 4;

QY 49 aacatcatcaatcaccctccatccgtgaacccgtgctgacgaatccaatccatccatccg 108
Db 49 ATCTGAAAAAGTCTACATTTCTGAACTTGACATCAACACATATTTATCTCCGACATC 108
QY 109 tctcgtacgcttccaaatcaaatcgcgttctcaagttcaactcgatccg---atcgac 165
Db 109 TCTGTTTCAACTCCCTGTTATACATATCCAGATGCTCAATGTTGGCGGCGATCAAC 168
QY 166 aagaatcaagatccagcgttcaatctggaatcttccaaatcgaagtatccggaagat 225
Db 166 GGCMAAGCTTACCACTCGTTAACACGAATCTTCAAGTTATTCGACAAAGGCCATG 228
QY 226 gcatcgtatatacaactctatgtaagaacttccaccctctcgtgacgtatcccg 285
Db 226 GACATGCAATCAACAGCATGTTCAACACTTACCGTTAGCTTGGCTGGCGGTTCCG 288
QY 286 aaatcactcaatccatc-----tctcgaacaatgaatcaacatcaatcgaatc 336
Db 286 AAAGTTCTGCTCTCCCACTGAAACAGTACGCGACATCACTATCATCATCATCT 348
QY 337 atggaatacaat-----tctggttgaagaatctctcgtgaactcgtgtgaa 381
Db 337 ATGAAGAACACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 408
QY 382 atcatcgtacgtcgcagacacacacacacacacacacacacacacacacacacacac 441
Db 382 CTGATCGACCTGTAAGAACTCCGCGGAGGAGTTGTCGATCACTTCCGCGACCTG 468
QY 442 cagatgatacaatcctcgtacatcaatcgtcgtgacgttctgttaccatccacaat 501
Db 442 CCGGACAAAGTTCAACGCGTACCTGCTTCAACAAATGGGTTTCACTATCACTAACGAT 528
QY 502 cgtcgtatatacccaaatctatacaagcgcgtcgtgacgaacgaacgaatcc 551
Db 502 CGTCTGCTTCTGCTACCTGTAATCAACGCGGTTGTAAGGCTTCCGCTGAAATCACT 568
QY 562 aatcgtgtaatacccaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 621
Db 562 GGTCTGGGCGCTATCCGCTGAGAACACACATCACTTAAAGTGGAGCTTGCAACAC 648
QY 622 actcaacgtacatctgatacaatctcaatctcgttgcagaagaactgaagaagaa 681
Db 622 GGTCTGGGCGCTATCCGCTGAGAACACACATCACTTAAAGTGGAGCTTGCAACAC 708
QY 682 gaaatcaagaactgtacgaacacagltccaaatctgtatcttcgaaagacttcg 741
Db 682 GAGATCGAAAAAAGTGTATACACTACTGCTATACACTTCTGCTGCTGCTGCTGCTG 768
QY 709 GAGATCGAAAAAAGTGTATACACTACTGCTATACACTTCTGCTGCTGCTGCTGCTG 768
Db 709 GAGATCGAAAAAAGTGTATACACTACTGCTATACACTTCTGCTGCTGCTGCTGCTG 768
QY 742 gactactgtgacgaacacacgttactatactgtcaatctgtacgaatccgaacaa 801
Db 742 GAGATCGAAAAAAGTGTATACACTACTGCTATACACTTCTGCTGCTGCTGCTGCTG 828
QY 802 gttgacgtcaacaatgtaggtatccgcggtgacgtgacgtgacgtgacgtgacgt 861

Db 829 GTTACGCTGAAAAA-----CATGACTGACTACATGTACTGACCAACGCCGCTCTAC 882
Qy 862 gtatgactccaacattaccctgaactctccctgtaacggtgtaccaaatcatc 921
Db 883 ACTACGGGTAACTAACATCTACTACTACGCGCTGTACCAACGGCTTAATCATCATC 942
Qy 922 aagaatacgcgtctgtaacaaagacaat 951
Db 943 AAACGCTACACTCCGACACACGAATGAT 972

RESULT 12
US-07-618-312A-1
; Sequence 1, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618.312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1359 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Clostridium tetani
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1356
; US-07-618-312A-1

Query Match 5.6%; Score 74.2; DB 1; Length 1359;
Best Local Similarity 47.3%; Pred. No. 3.9e-12;
Matches 223; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

Qy 349 tctggttgaaagtatactctgaactacggtgaaatcatctggaactctgcagacactcag 408
Db 376 TCTGTTGGAGTATGATCACTTAAGGTAATACTTAATATGAGCTTTAAAGATTCCCG 435
Qy 409 gaaatcaaacagcggtgtgtatataactctcagaatgatacatcatctggaactac 468
Db 436 GGAGAGGTAGACAATAATCACTTTAGGGATTTACCTGTGATTAATTAATGCTTATTAGCA 495
Qy 469 aatcgatgattctgtaccatcaccaacacgctggaatcaactccaatctacac 528
Db 496 AATTAATGGGTTTATTACTATTACTTAATGATTAATGCTTCTGCTAATTTGTATATA 555
Qy 529 aacggcgctcgtatcgaccagaacacgatactcaactctggtgaatccacacgcttcaat 588
Db 556 AATGAGTACTATTATGGGAAGTGCAGAAATTTACTGGTTTAGAGCTATTAGAGAGATAT 615
Qy 589 aacatcatgttcaaacctggaacggtgtgtgtaacactacacgctacatctgtatcaatc 648
Db 616 AATATAACATTTAAACTAGATGATGATTAATTAATTAATCAATAGCTTCTATTGATATA 675
Qy 649 tcaatctgtcgaacaaactggaacggaacgaacgaacgaacgaacgaacgaacgaacga 708
Db 676 TTTAGGATATTTTGCAGACATTAATCCAAAGAGATTGAAATTAATTAATCAAGTTAT 735
Qy 709 tccaattctgatactcgaagaactctggtgtgactacactgacgtacgacacaacgctac 768
Db 736 TTATCTATTAACCTTTTAAAGACCTCTCGGGAACCCCTTAGATATGATACAGATAT 795
Qy 769 tacatgctgaatctgtacgataccgaacaaatcagttgaactcaaatgta 819
Db 796 TATTTAATACGATAGCTTCTAGTTCTTAAGATGTTCAATTAATAATA 846

RESULT 13
US-08-110-786A-7
; Sequence 7, Application US/08110786A
; Patent No. 5443966
; GENERAL INFORMATION:
; APPLICANT: FAIRWEATHER, Neil Fraser
; APPLICANT: MAKOFF, Andrew Joseph
; TITLE OF INVENTION: Expression of tetanus toxin fragment C
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 5443966th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110.786A
; FILING DATE: 23-AUG-1993 1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777.337
; FILING DATE: 29-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00943
; FILING DATE: 20-JUN-1990
; APPLICATION DATA:
; APPLICATION NUMBER: GB 8914122.0
; FILING DATE: 20 June 1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000

Qy	769	tacatgctgaatctgttaacgataccgaacaaatacgttaacgta	819
Db	796	tatttaataaccagtagctctagttcttaagatgttcattgaaaaatata	846

RESULT 15

```

US-08-668-381A-6
Sequence 6, Application US/08668381A
Patent No. 5780024
GENERAL INFORMATION:
APPLICANT: Brown, Robert H.
APPLICANT: Fishman, Paul S.
APPLICANT: Francis, Jonathan W.
APPLICANT: Hosley, Betsy A.
TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TERANUS
TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
City: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/668,381A
FILING DATE: 21-JUN-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/000,473
FILING DATE: 23-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,164
REFERENCE/DOCKET NUMBER: 00786/269001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1858 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-668-381A-6

```

Query Match	5.6%	Score	74.2	DB 1	Length	1858
Best Local Similarity	47.3%	Pred. No.	4.6e-12			
Matches 223	Conservative	0	Mismatches	248	Indels	0
					Gaps	0

[illegible]

Db	1055	AAAGGAGTACTTAAAGGGAAGTGGAAAGAAATTA	CTGGTTT	AGGAGCATTA	AGAGAGGATTAAT	1114
Qy	589	aacatcatgttcaaacctgtgaacggtgttcgt	gtgacatccacgcgtacatcgtgatacaataac	648		
Db	1115	AAATTAACATTAAACACATGATACATGTAAATAA	ATAATTAACATACGTTCTTTATTAATA	1174		
Qy	649	tccaatctgttcgacaagaagcgtgaacgaataa	gaagccttgacgtgacacaacag	708		
Db	1175	TTTAGGATATTTTGCAGAAAGCATTAATCCAA	AGAGATTTGMAAAATTATACCAAGTTAT	1234		
Qy	709	tccaatctcgtglatccctgaagaagcttcctgg	gtgtgacactccgtgaagacaacacggtac	768		
Db	1235	TTATCTTTAAACCTTTTAAAGAGCTTCGGGGA	AAACCTTTACGATATGATACAGAAATAT	1294		
Qy	769	tacatgctgaatcgtgtacggtccgaacaana	taacgttggaagtcacaacatga	819		
Db	1295	TATTTAATACCAAGTACGCTTCTAGTTCTAA	AGATGTTTCAATTGAAAATATA	1345		

Search completed: September 16, 2002, 21:32:06
Job time: 15658 sec

.....

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 17:08:43 : Search time 4365.92 Seconds
(without alignments)
4117.791 Million cell updates/sec

Title: US-09-611-419A-1
Perfect score: 1332
Sequence: 1 gaattcgaacagctgcgtct.....aacgtcgcgtctaagaattc 1332

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_hnv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58.4	4.4	608	10	BI887904 ZF637-1-0
2	54.6	4.1	1101	12	CNS00100
3	53.6	4.0	537	10	BM277996 AS_tgz-51
4	53.6	4.0	551	10	BM278689 AS_tgz-66
5	53.2	4.0	394	9	AU060224
6	52.6	3.9	447	10	BI594803 AS_tgz-44
7	52	3.9	597	10	BM278181 AS_tgz-54
8	51.8	3.9	500	10	BM278805 AS_tgz-67
9	50.6	3.8	511	10	BM004465
10	50.4	3.8	569	10	BM278657 AS_tgz-65
11	50	3.8	641	10	BM181884
12	49.6	3.7	645	9	AI389106 CH20192.5
13	49.2	3.7	827	12	CNS02156
14	49	3.7	481	10	BF050030
15	48.8	3.7	628	9	AU060230
16	48.2	3.6	546	10	C22974
17	48.2	3.6	1101	12	CNS0100X

18	48	3.6	546	10	BF169335
19	48	3.6	556	10	BM277853
20	47.4	3.6	377	10	BI594814
21	47.4	3.6	450	9	AU060996
22	47	3.5	458	10	BI783045
23	47	3.5	518	10	BI782765
24	46.8	3.5	1175	10	BI872945
25	46.4	3.5	469	10	BF050033
26	45.4	3.4	386	10	BF050064
27	45.4	3.4	499	10	BM278692
28	45.4	3.4	536	10	BF050073
29	45.4	3.4	542	10	BM278786
30	45.4	3.4	554	10	C23753
31	45.2	3.4	501	10	BM277916
32	45.2	3.4	525	12	BM184460
33	45.2	3.4	525	12	CNS07PEA
34	45.2	3.4	1101	12	CNS000D1
35	45	3.4	564	10	BE060733
36	44.8	3.4	534	10	BM278087
37	44.6	3.3	1101	12	CNS016JY
38	44.4	3.3	522	10	BM278558
39	44.4	3.3	681	12	CNS02EOD
40	44.2	3.3	325	12	CNS03PK8
41	44.2	3.3	429	9	AU052930
42	44.2	3.3	500	12	BE7199
43	44.2	3.3	519	10	BM278164
44	43.8	3.3	535	10	BI594948
45	43.6	3.3	443	10	BM039802

ALIGNMENTS

RESULT 1

BI887904

LOCUS ZF637-1-002159 zebrafish shield stage whole embryo cDNA library
DEFINITION MPMGP637 Danio rerio cDNA clone MPMGP637_10F4:MPMGP637F0410 5',
mRNA sequence.

ACCESSION BI887904 608 bp mRNA EST 12-OCT-2001
VERSION BI887904.1 GI:16095175

KEYWORDS zebrafish.
SOURCE EST.
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 608)
AUTHORS Clark,M., Aanstad,P., Hennig,S., Johnson,S.L. and Lehrach,H.
TITLE EST sequencing of a zebrafish shield stage cDNA library normalised
by oligonucleotide fingerprinting

JOURNAL Unpublished (2001)
COMMENT Contact: Hennig S
Laboratory 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inmestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
5' EST sequencing of clones from a zebrafish shield stage library,
normalised from 55,000 starting clones by oligonucleotide
fingerprinting
High quality sequence stop: 608.
location/Qualifiers

FEATURES
source 1..608
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="MPMGP637_10F4:MPMGP637F0410"
/clone_lib="zebrafish shield stage whole embryo cDNA
library MPMGP637"
/tissue_type="whole embryo"
/dev_stage="shield stage, 6 hrs post-fertilisation"
/lab_host="E.coli, XLI blue MRF"

```

/notice=Vector: site1:NotI; site2: SalI;
oligo-dr-NotI primed, SalI adaptors, directionally cloned
library normalised by oligonucleotide fingerprinting"
BASE COUNT      207 a      218 c      41 g      140 t
ORIGIN

```

Query Match	4.4%	Score 58.4	DB 10	Length 608
Best Local Similarity	49.1%	Pred. NO. 0.00034		
Matches 181, Conservative	0	Mismatches 187	Indels 1	Gaps 1

Oy 448 atcaacatctctgactatacatcaa-tcgctgagatcttgittaccatcaccaacaatgcgtc 506
| | | | | | | | | | | | | | | | | |
Db 5 ATCAACTTCTTCAGGCTACATCAACTNNCTCCAGCTACATCAACTTCTTCACAACAACAA 64

Qy 507 gaataactccaanaatcatcatcaacgycgctcgtatcgaccagaaacgatactccaatct 566
 || ||| | ||||| | ||| ||| |||||
 Db 65 CACCAACTTCTCCAGTACATCAACTTCTCCAAACAACAACAACAACAACAATCTTCCAGCTA 124

Oy 567 gggaacatccagcgttctaatacatcatgtccaactgagcgttgctcgtagaacata 626
||| || | | | ||| || | | | ||| |
Db 125 catcaactttccagctacatcaactttccagctacatcaactttccagctacacaa 184

Qy 627 ccgctacatctgtaacaaataactcaatctgttcgacaaagaactgaacgaaanaaat 686
 185 CTTCCTCAGCTACATCAACTTCTCCCAACACACACACAACTTCTCCACACACACAA 244

OY		687	caagacacctgtacgacaaccacgtgccaatctcgtgatccatgaagaacttcggggtgcta	746
Db		245	CAACAACTTTCCTCAGCTATCATCAACTTTCTCCACAAMCACAAACAACTTCTCCAGCTA	304

Qy 747 ccgcagtagcacaacccgtactacatgctgaatctgtacgatccgaacaaatacgltta 806
- - - - -
Db 305 CATTCACTTCTCCAAATCATCAACACAAGTTCTTCAGGTACTACATCAACAMCAAGTTCTTCAG 364

Qy	807	cgtcacaca	815
Db	365	CTACACAA	373

RESULT 2
CNS00LO0

LOCUS	GENOME	COORDINATES	REFERENCE
Drosophila melanogaster genome survey sequence TET3 end of BAC Drosophila	Genome Survey Sequence	1000000-1000000	TET3 end of BAC Drosophila
BACR32d23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	Genomic Survey Sequence	1000000-1000000	BACR32d23 of RPCI-98 library from Drosophila melanogaster (fruit fly)

ACCESSION	AL068807
VERSION	AL068807.1
KEYWORDS	GSS.
SOURCE	fruit fly.

ORGANISM *Drosophila melanogaster*
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

JOURNAL TITLE
AUTHORS
REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequenc

BP 191 91006 EVRY cedex FRANCE (E-mail : segref@genoscope.cnrs.fr)
- Web : www.genoscope.cnrs.fr)
Determination of this BAC-end sequence was carried out as part of a collaborative project between the French Genomic Institute (Genoscope) and the French National Institute for Research in Agriculture (INRA) in collaboration with the Barkeley Drosophila Genome Project (BPGP).

The BGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BGP *Drosophila melanogaster* BAC library was created by Genentech, Inc.

Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

isogenic strain Y2: cn bw sp, the same strain used for the BDG pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or

```

filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.bufileo.edu/drosophila\_bac.htm.
location/Qualifiers
1. 1101
FEATURES
source
```

BASE COUNT	155 a	166 c	7 g	284 t	489 others
------------	-------	-------	-----	-------	------------

Query Match	Score	DB	Length
4.18;	54.6;	DB 12;	length 1101;

Matches 100; Conservative 219; Mismatches 245; Indels 3; Gaps 25

tctaccttaactgatatcatcaagaacatcatcatacttcacatccctgaacctgcgtac 84

Db 498 TCTTTCACATCATCMMHNCSSMHHNMACTMMIMKUYAAVTCSTHYMCNHMSYTHCSMTPLAM 557

Qy 85 gaatccaataaccctgatacgcactgtctcgctacagcttccaaaatacaaatcggtttcta aa 144

Db 558 CMMTCATATCMTCMHNAYCMYCCMCYTTCCMCAHHTMANACAMTMNAHCHTHTTMC 617

Db 618 MTMTTMAATTCMCAHTMAHTMTMAMCAAHNNITWTMMAMHMAIACHACNNITTCAMM 677

Db 678 MMKALMCHHHMMMAHHTALMTCAHMAAAMAMHTMTTW---AMMACHRAMMMCC 734

Db 735 CSCAMCMAMAMAHMAMCMAMHHCMMCMCMAMMCMCMAMAMTACACACAAAMAAAMTATM 794

QY 325 atcatcaactgcataqaaaacaattctgtgttgaagaatattctctaactacgttgaatc 384

Db 795 MHCAMMMMYCMMMAHCTMTACCMCMNCYCACMTCTCCCMCMMAMTMMAMMTTTH 854

Db 855 CHTTTCMHHNTTTCMTMCCSMMMMMAHNMHCMMHTHHMHNAHNTMCMCCCTTMM 914

Ov 445 atgataacatctctgaactacatcaatcgcgtatgacatcttcgattacatcaaccaaatcgt 504

Db 915 WHAHTWVHHVTHHHNHWCAAMMMACHAMMMSCMMWVSCMMSCMMVTVHNAAMNHT 974

0v	565	ctgaatacattccacgccttctaataac	591
Db	975	TTTTTMTTCMYCMAMGMAMAMAMCMMAAMATCMAA	1034

```

:      | | : : |      | : : : : |
Db  1035 МАМАААААММММММССУМСАМТМАММННАУ 1061

```

RESULT	3
BM277996	
LOCUS	537 bp
DEFINITION	mRNA
Accession No.	U08453
Source	Homo sapiens adult male testis normal
Size	537 bp
Molecule type	linear
Date Recd.	EST 20-DEC-2000

Accession	Version	Accession	Version	Accession	Version	Accession	Version
Alan Scott Ascaris suum cDNA clone AS_tgz_51F02 5', mRNA sequence							
BM277996							
BM277996.1	GI:17971254						

SOURCE	ORGANISM
pig roundworm.	<i>Ascaris suum</i>
	Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascarididae

REFERENCE
AUTHORS

1 (bases 1 to 537)
Blaxter, M. L., Parkinson, J., Whitton, C., Daub, J., Guiliiano, D., Ha
, N., Quayle, M. and Barrell, B.

TITLE Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk

The library was prepared by Michelle Lizotte-Waniewski for Alan Scott, Johns Hopkins University Medical School, Baltimore. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).

PCR Primers
FORWARD: T3
BACKWARD: T7PL
Plate: 51 row: F column: 02
Seq primer: SKPL
High quality sequence stop: 478.
Location/Qualifiers
1. 537
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_lib="As_tgz_51F02"
/clone_lib="Ascaris suum adult male testis germinal zone from Alan Scott"
/sex="Male"
/dev_stage="Adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI; Library was made from dissected testis germinal zone from adult male Ascaris suum collected from abattoirs. Constructed by Michelle Lizotte-Waniewski for Alan Scott, Johns Hopkins University Medical School, Baltimore, MD."

BASE COUNT 193 a 149 c 72 g 123 t
ORIGIN

Query Match 4.0%; Score 53.6; DB 10; Length 537;
Best Local Similarity 48.6%; Pred. No. 0.0055;
Matches 180; Conservative 0; Mismatches 184; Indels 6; Gaps 1;

QY 368 tgaactcgtgaaatcgtgactctgcaagacactcaggaatcaaacagcgtgtg 427
DB 49 TGTACAACTTACAACTGTCACAAAGAACCCACTGACAACTTCAACAAAGCAACTTTTAA 108
QY 428 tatcaaatcactcagatgatcaaatcctcgtactcaatcgaatcgtgactcgtta 487
DB 109 GCAACTTCAACAGTAACCTTCAACAGCCATTTCAACAGCAACTTCAACAACTTCAATA 168
QY 488 ccatcacaacacatcgtctgaaataactcoaaatctacatacagcggtctgacgacc 547
DB 169 ACAACTTACGACACTCTGTAACAACTTCAATTTTTCGCAACTCTGAGACACG 228
QY 548 agaaacgcatcctcaatctgggttaacatccacgctcttaatacatalcgttcaaatgg 607
DB 229 GGAG-----CTACAACTTGAACAACCTCAACACACTCAATTAACAACGTAATCTGA 282
QY 608 acggtgtcgtgacactcagcgtacatctggtacaataactcaatctgltcgacaag 667
DB 283 ACAACAACTTCAAGCAATTTCAACAATACTTCAACAACTTGAACAGCAACTTCAACA 342
QY 668 aactgaagaaaagaataaagaacgtgtgacgaacacagtcgaattctgtatcctga 727
DB 343 ACAGTAACCTTCAACAACCTTCAACAGCAACAACTTCAAGCAACTTCAACAACAGTAACCTTCA 402
QY 728 aagactctg 737
DB 403 ACAACAGCAG 412

RESULT 4
BM278689

LOCUS BM278689 551 bp mRNA linear EST 20-DEC-2001
DEFINITION As_tgz_66A05_SKPL Ascaris suum adult male testis germinal zone from Alan Scott
ACCESSION BM278689
VERSION BM278689.1 GI:17971947
KEYWORDS EST.
SOURCE pig roundworm.
ORGANISM Ascaris suum
; Ascarididae; Ascaris.
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascarididae

REFERENCE 1 (bases 1 to 551)
AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guillano,D., Hall,N., Quayle,M. and Barrell,B.
TITLE Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk

The library was prepared by Michelle Lizotte-Waniewski for Alan Scott, Johns Hopkins University Medical School, Baltimore. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).

PCR Primers
FORWARD: T3
BACKWARD: T7PL
Plate: 66 row: A column: 05
Seq primer: SKPL
High quality sequence stop: 510.
Location/Qualifiers
1. 551
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_lib="As_tgz_66A05"
/clone_lib="Ascaris suum adult male testis germinal zone from Alan Scott"
/sex="Male"
/dev_stage="Adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI; Library was made from dissected testis germinal zone from adult male Ascaris suum collected from abattoirs. Constructed by Michelle Lizotte-Waniewski for Alan Scott, Johns Hopkins University Medical School, Baltimore, MD."

BASE COUNT 208 a 160 c 64 g 119 t
ORIGIN

Query Match 4.0%; Score 53.6; DB 10; Length 551;
Best Local Similarity 49.6%; Pred. No. 0.0055;
Matches 137; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 437 acctcagaatgatacaatcctcgtactacataatcgtcgatctcgttaccatca 496
DB 102 ACAGTAACCTTCAACAACAGCAATTTCAACAGCAACTCTGCAACTTCAATTAACACCTTA 161
QY 497 acaatcgtctgaaataaataatcacaatacgaagcgtctgactgacgaaacgga 556
DB 162 GCAACTCTGTAACAACTTCAATTAACAACTCTGCAACCTCTGAGCAACCTCACTTACA 221
QY 557 tctcaatctgggttaacatccacgctcttaataacatcgttcaaacctgacggtgtc 616
DB 222 ACTTGAAACAACTTCAACAACCTTCAACAACTCTCAATTAATTAACAACTTCAACA 281
QY 617 gtgacactcaccgctacacatctgataatataatctggttgcagaagaacggaag 676
DB 282 GTCTCAACATTAACCTTCAACAACCTTGAACAGCAAGCAATTTCAACAGCAATTAACCTTCAACA 341
QY 677 aaaaagaataaagacctgtacgacaacggtcca 712

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota: Metazoa: Nematoda; Chromodorea; Ascaridida; Ascaridoidea ; Ascaridae; Ascaris. 1 (bases 1 to 447)	Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guilleman, D., Hall , N., Quayle, M. and Barrell, B.	Edinburgh University/Sanger Centre Nematode EST Project Unpublished (2000)		
	Contact: Blaxter ML Institute of Cell, Animal and Population Biology University of Edinburgh Asmworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK. Tel: +44 131 650 6760 Fax: +44 131 670 5450 Email: mark.blaxter@ed.ac.uk			
	The library was prepared by Michelle Lizotte-Waniewski for Alan Scott, Johns Hopkins University Medical School, Baltimore. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The sequence contained a Polya tail (trimmed)			
	PCR primers FORWARD: T3 BACKWARD: T7PL Plate: 44 row: C column: 11 Seq primer: SKPL High quality sequence stop: 447. Location/Qualifiers 1..447			

[illegible]

ACCESSION	Atan.Scott	Ascaris	sum	CDNA	clone	As_tgz_54C05_5',	mRNA	sequence.
VERSION	BM278181							
KEYWORDS	BM278181.1	GI:17971439						
SOURCE	EST.							
ORGANISM	pig roundworm.							
	Ascaris suum							
	Eukaryote; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea							
	; Ascaridae; Ascaris.							
REFERENCE	1 (bases 1 to 597)							
AUTHORS	Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Gulliano,D., Hall							
	N., Quayle,M. and Barrell,B.							
TITLE	Edinburgh University/Sanger Centre Nematode EST Project							
JOURNAL	Unpublished (2000)							
COMMENT	Contact: Blaxter ML							

University of Edinburgh
Ashworth labs, King's Buildings, West Main Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared by Michelle Lizotte-Waniewski for Alan
Scott, Johns Hopkins University Medical School, Baltimore.
Sequencing was performed by the Pathogen Sequencing Unit, Sanger
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).
PCR primers
FORWARD: T3
BACKWARD: T7PL
Plate: 54 row: C column: 05
Seq primer: SKPL
High quality sequence stop: 502.

FEATURES	Location/Qualifiers
source	1. .597

```

/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone="As_tgz_54C05"
/clone_1ib="Ascaris suum adult male testis germinal zone
from Alan Scott"
/dev_stage="Adult"
/sex="Male"
/notes="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
library was made from dissected testis germinal zone from
adult male Ascaris suum collected from abattoirs.
Constructed by Michelle LiCotte-Waniewski for Alan Scott,
Johns Hopkins University Medical School, Baltimore, MD."

```

Query Match	3.9%	Score 52	DB 10	Length 597
Best Local Similarity	49.3%	Pred. NO.	0.015	
Matches 136	Conservative	0	Mismatches 140	Indels 0
				Gaps 0

Qy	437	acttcagatatacaacatctctgactcaatccaatcgcttgatctctgtttccatcacca	496
Db	141	ACAATACTTCAACACAGCAATTTTCAACACAGCAACTGAGCACTTCATTAACAACTTA	200
Qy	497	acaatcgtctgtaataactccaanaatctataccaagcgcgtctgatacgccaagaaccca	556
Db	201	GCAACTCTCTGTAACAACTTCATTAACAACCTTCAGCAACCTGTCTGAGCAAACTCAATTTCA	260
Qy	557	tctccaatctggtcaacatccagcgtctctataatacatgatgttcaaatgtgaagttctc	616
Db	261	ACTTGMAACAACCTCAACAACCTCAATTAACAACCTCAATCTCAACAACAATCTCAACA	320
Qy	617	gtgaacactaccgcgtacatctgataccaatactccaatctgttcgcgaacaagaactgaa	676
Db	321	ATTCACAACTTAACCTCAACAACCTTACACAGACAACCTTAACAAGCAGTAATCTCAACA	380
Qy	677	aaaaagaatcaaaagacctgtagagacaacagctcca	712
Db	381	ACAGTAACCTTCAGCAACTTCAACAGCAACAACCTTCA	416

RESULT	8
LOCUS	BM278805
DEFINITION	BM278805 500 bp mRNA linear EST 20-DEC-2001 As_tg_67Ff10.SKPL Ascaris suum adult male testis germinal zone from Alan Scott Ascaris suum cDNA clone As_tg_67Ff10 5', mRNA sequence.
ACCESSION	BM278805
VERSION	BM278805.1 GI:17972063
KEYWORDS	EST.
SOURCE	p19 roundworm.
ORGANISM	Ascaris suum

REFERENCE	1 (bases 1 to 500)
AUTHORS	Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Gulliano, D., Hall, N., Quayle, M. and Barrell, B.
TITLE	Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL	Unpublished (2000)
COMMENT	Contact: Blaxter ML

Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared by Michelle Lizotte-Maniewski for Alan
Scott, Johns Hopkins University Medical School, Baltimore.
Sequencing was performed by the Pathogen Sequencing Unit, Sanger
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The
sequence contained a Polya tail (trimmed)

```

BACKWARD: T7PL
plate: 67  row: F  column: 10
Seq primer: SKPL
High quality sequence stop: 500.
      location/Qualifiers
1. .500
FEATURES
source
```

```

/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone="As taz_67F10"
/clone_lib="Ascaris suum adult male testis germinal zone
from Alan Scott"
/sex="Male"
/dev_stage="Adult"
/notes="Vector: Lambda zap II; Site_1: EcoRI; Site_2: XhoI;
library was made from dissected testis germinal zone from
adult male Ascaris suum collected from abattoirs.
Constructed by Michelle Lisotte-Waniewski for Alan Scott,
Johns Hopkins University Medical School, Baltimore, MD."

```

Query Match	3.9%	Score 51.8;	DB 10;	Length 500;
Best Local Similarity	50.8%;	Pred. 0.015;		
Matches 153;	Conservative 0;	Mismatches 142;	Indels 6;	Gaps 1;

QY	437	actctcaga	tgatacaacatctctctgacatc	atcaatcgtgtgatcttg	tccatcacca	496
Db	70	ACAGTAACCTTCAACCAACGCAATTTCAACACACACCTTCAGAACTTCATATAACCAACTTAA				129
QY	497	acaatcgcgtgtaataactcccaaaalctatcatccaagcgctctgatacgacacagaaaccca				556
Db	130	GCAACTCTCTGACAACTTCATATAAACCCCTCAGCAACTCTCTGAGCAACCTCAAG----				184
QY	557	tctcaaatcttggtgaactctcagcgcttctaataacatcatgttcaaatgtagcgtgtgc				616
Db	185	-CTTAACTTTGACACCACTCAACACCTTCATATAACCTTAACATATCTCAACACCAATTC				243
QY	617	gtgaacactcacgcgtacatactgataccaacttaacttaactcgtgtcgacaaagaactgaagc				676

Db 244 TCAGCAATCTGAACAATAAACCCTGACAAACACTTAGAACAGCAACTTTCACAAACAAGAATTAACCT 303
Oy 677 aaagaataacaagacctgtacgaaccaggcccaattcctggtatcctgaagacttct 736
 | | | | | | | | | | | | | | | | | |
Db 304 TCCAAACCTTGCAACGACAACAACCTTCACCAACCTTCACAAACAGTAATTCAACAACAGCAGCA 363
Oy 737 g 737
 |
Db 364 G 364

RESULT 9
BM004465 511 bp mRNA linear EST 25-OCT-2001
LOCUS BM004465
DEFINITION tgstzya86f03.y1 tgyeg. Partially sporulated oocyst cDNA Toxoplasma gondii cDNA clone TGSTZYA86F03.Y1 5' mRNA sequence.
ACCESSION BM004465
VERSION BM004465.1 GI:16439239
KEYWORDS EST.

SOURCE Toxoplasma gondii.
ORGANISM Toxoplasma gondii.
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
1 (bases 1 to 511)
Tang,R., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,
E., Bennett,J., Franklin,C., Tsagarisvilli,R., Ronko,I., Kennedy,
S., Maguire,L., Waterston,R. and Wilson,R.
Toxoplasma EST Project
Unpublished (2001)

TITLE Toxoplasma EST Project
JOURNAL Contact: Clifton, S.
COMMENT Toxoplasma EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: toxo@watson.wustl.edu
Contact David Sibley (toxosteborcim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
High quality sequence stop: 433.
Location/Qualifiers
1..511
/organism="Toxoplasma gondii"
/strain="VEG (Type III)"
/db_xref="taxon:5811"
/clone="TGSTZYA86F03.Y1"
/clone_lib="TGYEG partially sporulated oocyst cDNA"
/dev_stage="Partially sporulated oocysts"
/note="Vector: Modified pluescript (PBSSK+); Site_1:
BamHI; Site_2: EcoRI; Polya mRNA from partially-sporulated
oocysts was converted to cDNA using the
template-switching PCR method (SMART cDNA, Clontech Inc.,)) and
size selected on SizeSep 400 columns (Amersham
Pharmacia Biotech Inc.). First strand was reverse
transcribed using the CPS III-oligo-dT primer and a 5'
template switch primer (Smart IV primer). The product of
the first strand synthesis was PCR amplified using the
same primer set and the fragments were digested with SfiI.
The fragments were size selected, ligated into a modified
pluescript vector containing directional flit sites, and
electroporated into DH10B or DH12S cells. Vector: Sfil
sites were added to the multiple cloning region of
pbuescript SK+ between the BamHI/EcoRI sites. The
modified polylinker has the following sequence:
5'-GAATTCGGGCACATTACGCC(G)n-- insert -
GGCGCCCTCGGCCACGATCC3'where n=3-4 G nucleotides. Library
Source: Michael White, Maria E. Jerome, Emily A.Johnson,
Jay A. Radke, Montana State University, Clone
Availability: David Sibley, Washington University"

FEATURES
Source

BASE COUNT 213 a 181 c 61 g 56 t

Query Match	Similarity	3.8%	Score 50.6	DB 10	Length 511
Best Local	Similarity 46.9%	Pred. No. 0.031			
Matches 158	Conservative 0	Mismatches 179	Indels 0	Gaps 0	
QY	432	caaatcactcctagatgatcaacaatctctcgtactacatcgaatcgtctgtagcttcgtaccat	491		
DB	169	CTACTACTACAACTACGACACCAACCAACTACACAACTACACAACTACACAACTACACAA	228		
QY	492	caccacaatcgtctcgtaatcctccaatctacatcaagaagccggtctcgtacgcacgaa	551		
DB	229	CAACGACCAACCAACCAACCAACTACACCAACCAACCAACCAACCAACCAACTACAC	288		
QY	552	accgactcccaatcctcgttgtaatccacgcctcctaataacatcatcgttcaacatcgagcg	611		
DB	289	CAACGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACTACGA	348		
QY	612	ttgtcgtgacactcaccgcgtcatcatctgtagcaaatcattcaatctgttcgcaagaact	671		
DB	349	CTACGACCAACTACACTACACTACCAACTACCAACGACCAACGACCAACCAACTACAA	408		
QY	672	gaacgaaaaagaatccaagacctgttagacacacgaatcccaatctgttatccctgaaga	731		
DB	409	CAACGACCAACCAACCAACCAACCAACTACCAACGACCAACGACCAACCAACTACGA	468		
QY	732	ctctcgtggtgactaccctgcgcgtacgacgaacacgctac	768		
DB	469	CTACGACCAACTACACTACGACCAACTACCAACGACCAACTACCAACTACCAACTAC	505		
RESULT 10					
LOCUS	BM278657				
DEFINITION	As_tgz_65E06_SRP Ascaris suum adult male testis germinal zone from Alan Scott Ascaris suum cDNA clone As_tgz_65E06_5', mRNA sequence.				
ACCESSION	BM278657				
VERSION	BM278657.1				
KEYWORDS	EST.				
SOURCE	pig roundworm.				
ORGANISM	Ascaris suum				
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea				
AUTHORS	1 (bases 1 to 569)				
TITLE	Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Gulliano,D., Hall				
JOURNAL	,N., Quayle,M. and Barrell,B.				
COMMENT	Edinburgh university/Sanger Centre Nematode EST Project , Unpublished (2000) Contact: Blaxter ML Institute of Cell, Animal and Population Biology University of Edinburgh Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JY, UK. Tel: +44 131 650 6760 Fax: +44 131 670 5450 Email: mark.blaxter@ed.ac.uk The library was prepared by Michelle Lizotte-Waniewski for Alan Scott, Johns Hopkins University Medical School, Baltimore. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The sequence contained a polyA tail (trimmed) PCR primers FORWARD: T3 BACKWARD: T7/PL Plate: 65 row: E column: 06 Seq primer: SKPL High quality sequence stop: 518. Location/Qualifiers 1..569 /organism="Ascaris suum" /db_xref="taxon:6253" /clone="As_tgz_65E06" /clone_lib="Ascaris suum adult male testis germinal zone from Alan Scott"				


```

/sex="Male"
/dev_stage="Adult"
/note="Vector: Lambda Zap II; Site.1: EcoRI; Site.2: XhoI;
Library was made from dissected testis germinal zone from
adult male Ascaris suum collected from abattoirs.
Constructed by Michelle Lizotte-Waniewski for Alan Scott,
Johns Hopkins University Medical School, Baltimore, MD."
BASE COUNT      213 a      164 c      69 g      123 t
ORIGIN

```

```

Query Match      3.8%; Score 50.4; DB 10; Length 569;
Best Local Similarity 48.9%; Pred. No. 0.037;
Matches 135; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 437 actctcagatgatacaacatctctgactacatcgcgtgactctgttaccaccca 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 119 ACAGTAACCTTCACACAGCAATTTCAACAGCAACCTTCAGCAATTAACAACCTTA 178
QY 497 acaatcgtctgaataactccaaatctatcaacgagcgctctgtagcagaaccca 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 179 GCAACTCTCTCAACACTTCAATTAACACCTCAGCAACTCTCTGAGCAACTCAGCTACA 238
QY 557 tctccaatctggttaacatccacgcttcttaataatcatgttcaactggagcgtgtc 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 239 ACTTGAAACACCTTCACACACACTTCATACACACTTCATCTCAACAACATCTCAGCA 298
QY 617 gtgacactcaccgctacatctgtagcaatactcaatctgttggacaagaactgaacg 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 299 ATCTCAACAAATACCTTCACACACTTACAGCAGCAACTTCACAGCAAGTAACCTTCAACA 358
QY 677 aaaaagaatcaagaacctgtgacgacaccagttcca 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 359 ACAGTAACCTTCAGCAACTTCAACAGCAACAACTTCA 394

```

```

RESULT 11      641 bp      mRNA      linear      EST 11-DEC-2001
BM181884
LOCUS
DEFINITION
f51b11.y1 Sugano STD adult male Dantio rerio cDNA clone 5412044 5'
similar to contains element TARI repetitive element ;, mRNA
sequence.
ACCESSION
BM181884
VERSION
BM181884.1 GI:17512842
KEYWORDS
EST.
SOURCE
zebrafish.
ORGANISM
Dantio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Dantio.
1 (bases 1 to 641)
REFERENCE
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,A., Hillier,L., Kucaba,T., Martin,D., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Pearson,B.,
Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shu,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
AUTHORS
Masnu Zebrafish EST Project 1998
JOURNAL
Unpublished (1998)
COMMENT
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zdraifshewatson.wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution information can be found through the I.M.A.G.E.
Consortium/LINL send email to: infoimage.linl.gov
Seq primer: T3 ET from Amersham
High quality sequence stop: 527.
Location/Qualifiers
1. 641

```

```

/organism="Dantio rerio"
/db_xref="taxon:7955"
/clone="5412044"
/clone_11b="Sugano STD adult male"
/sex="Male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="Vector: pME18S-FL3; Site.1: DraIII (CACCATGNG);
Site.2: DraIII (CACCATGNG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[TTGTGGCCCTTACG6], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGNG, 3' site
CACCATGNG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAGCTGCG and 3' end primer
CGACCTGCGAGCTCGAGCA."
BASE COUNT      218 a      234 c      37 g      152 t
ORIGIN

```

```

Query Match      3.8%; Score 50; DB 10; Length 641;
Best Local Similarity 46.3%; Pred. No. 0.049;
Matches 201; Conservative 0; Mismatches 230; Indels 3; Gaps 1;

```

```

QY 235 tacaactctatgtacgaanaactctccacactctctctgtagcgtatcccgaaatactc 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 175 TACATCAACTTCTCCACACAAACAACACTTCTCCACACAAACAACACTTCTCCAGCTACGC 234
QY 295 aactcattctctgtgacaaatgaataaccatcatcaactgcaatggaacaa---ttct 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 235 AACCTTCACACAAACAACAACTTCTCCAGCTACATCACTTCACACAAACAACAACTTCTCC 294
QY 352 ggttggaagaatctctgtagcagtggaatcatctgtagccttgaggagacacagga 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 295 ACTTACAACTTCTCCAGCAACAACATCACTTCTCCAGCTACATCACTTCTCCAGCAAC 354
QY 412 atcaaacagcgtgtgtatcaataactctcagatgatcaaatctctgatacatcaat 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 355 AACCAACTTCTCCAGCTACATCACTTCTCCAGCAACAACAACTTCTCCAGCTACATCAAC 414
QY 472 cgtgtgattctgttaccatcaacaaatcgtctggaataaactccaaatctacataac 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 415 TTCTCCACAAACAACAACTTCTCCAGCTACATCACTTCTCCAGCAACAACAACTTCTCC 474
QY 532 ggcgcgtctgtagcagcagaacacgactcgaatctggttaacatccagctctataac 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 475 AGCTACATCACTTCCACACAAACAACACTTCTCCAGCTACATCACTTCTCCAGCAACAAC 534
QY 592 atcatgttcaaaactgagcgtgtgtctgtagcactcagctacatctgatacaatactc 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 535 AACTTCTCCAGCTACATCACTTCAACACAAACAACAACTTCTCCAGCTACATCACTTCC 594
QY 652 aactcgttgcagaa 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 595 AACCAACACAAACA 608

```

```

RESULT 12      645 bp      mRNA      linear      EST 23-APR-2001
AT389106
LOCUS
DEFINITION
GH20192.5prime GH Drosophila melanogaster head P072 Drosophila
melanogaster cDNA clone GH20192 5, mRNA sequence.
ACCESSION
AT389106
VERSION
AT389106.2 GI:13758704
KEYWORDS
EST.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:29:41 : Search time 7254.67 Seconds
(without alignments)
3816.274 Million cell updates/sec

Title: US-09-611-419A-3
Perfect score: 1323
Sequence: 1 gaattcgaaacgctgtctac.....aacgtccgctgaagaattc 1323

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb-ba:*
2: gb-bhg:*
3: gb-ln:*
4: gb-om:*
5: gb-ov:*
6: gb-pat:*
7: gb-ph:*
8: gb-pl:*
9: gb-pr:*
10: gb-ro:*
11: gb-sts:*
12: gb-sy:*
13: gb-un:*
14: gb-vl:*
15: em-ba:*
16: em-fun:*
17: em-hum:*
18: em-ln:*
19: em-mu:*
20: em-om:*
21: em-or:*
22: em-ov:*
23: em-pat:*
24: em-ph:*
25: em-pl:*
26: em-ro:*
27: em-sts:*
28: em-un:*
29: em-vl:*
30: em-htg-hum:*
31: em-htg-inv:*
32: em-htg-other:*
33: em-htg-inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query Match Length DB ID Description

1	1304.4	98.6	1330	6	AR000029	AR000029 Sequence
2	1304.4	98.6	1330	6	AR169140	AR169140 Sequence
3	1304.4	98.6	1330	6	AX036243	AX036243 Sequence
4	1304.4	98.6	1338	12	XXU22962	U22962 Synthetic b
5	1304.4	98.6	1402	6	AR000030	AR000030 Sequence
6	1304.4	98.6	1402	6	AR169141	AR169141 Sequence
7	1304.4	98.6	1402	6	AX036246	AX036246 Sequence
8	727.6	55.0	4292	1	CBNT0A	X52068 C.botulinum
9	727.6	55.0	4835	1	CBNT0A	M30196 C.botulinum
10	726.4	54.9	3891	6	AR000031	AR000031 Sequence
11	726.4	54.9	3891	6	AR169142	AR169142 Sequence
12	726.4	54.9	3891	6	AX036248	AX036248 Sequence
13	662	50.0	4067	1	CBNT0A	X73423 C.botulinum
14	423.2	32.0	1299	12	AF251281	AF251281 Synthetic
15	421.6	31.9	1313	6	A58946	A58946 Sequence 6
16	228	17.2	4073	1	CBNT0A	X68262 C.botulinum
17	227	17.2	1293	6	A58945	A58945 Sequence 5
18	227	17.2	4199	1	CBNT0A	L35496 Clostridium
19	227	17.2	4209	1	CBNT0A	X81714 C.botulinum
20	195.8	14.8	3835	1	AB037705	AB037705 Clostridi
21	195.8	14.8	3835	1	AB037706	AB037706 Clostridi
22	195.8	14.8	3835	1	AB037707	AB037707 Clostridi
23	195.8	14.8	3835	1	AB037708	AB037708 Clostridi
24	195.8	14.8	3835	1	AB037709	AB037709 Clostridi
25	195.8	14.8	3835	1	AB037710	AB037710 Clostridi
26	195.8	14.8	3835	1	AB037711	AB037711 Clostridi
27	195.8	14.8	3835	1	AB037712	AB037712 Clostridi
28	195.8	14.8	3835	1	AB037713	AB037713 Clostridi
29	195.8	14.8	3835	1	AB037714	AB037714 Clostridi
30	195.8	14.8	4030	1	CBNT0A	X62083 C.botulinum
31	195.8	14.8	4030	1	CBNT0A	X62083 C.botulinum
32	194.2	14.7	4017	6	AX088262	AX088262 Sequence
33	194.2	14.7	4017	6	AB039264	AB039264 Clostridi
34	191	14.4	3849	1	CBNT0A	X62088 C.botulinum
35	191	14.4	4234	1	CBNT0A	M92906 Clostridium
36	180.8	13.7	9325	1	CBNT0A	Y13631 Clostridium
37	180.8	13.7	9325	1	CBNT0A	I28431 Sequence 3
38	166.4	12.6	1359	6	A49987	A49987 Sequence 4
39	166.4	12.6	3712	6	A49987	A49987 Sequence 4
40	166.4	12.6	3754	6	A37074	A37074 Sequence 17
41	166.4	12.6	3754	6	A42478	A42478 Sequence 6
42	166.4	12.6	3769	6	A37075	A37075 Sequence 18
43	166.4	12.6	3769	6	A49988	A49988 Sequence 5
44	166.4	12.6	4366	6	A42484	A42484 Sequence 12
45	166.4	12.6	4378	6	A42481	A42481 Sequence 9

ALIGNMENTS

RESULT 1
LOCUS AR000029 1330 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5736139.
ACCESSION AR000029
VERSION AR000029.1 GI:3962560
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1330)
AUTHORS Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.
TITLE Treatment of Clostridium difficile induced disease
JOURNAL Patent: US 5736139-A 22 07-APR-1998;
FEATURES
source location/Qualifiers
1..1330
BASE COUNT 400 a 339 c 246 g 345 t
ORIGIN

Query Match 98.6%; Score 1304.4; DB 6; Length 1330;
Best Local Similarity 99.9%; Pred. No. 0;

SFQID 3

	Matches	1305, Conservative	0, Mismatches	1, Indels	0, Gaps	0
OY	12	gatgtcacccttactgaatacatacaagaacatcaatcatccatccatcgtgaacctgcg	71			
Db	12	gctgtcttacccttccactgataatcatcaagaacatcatcatatcctcatcctcaacctgcg	71			
OY	72	ctacgaatccaatacaaccgtgatccgactctctgcgcacgcttccaaatcaacatcggttc	131			
Db	72	ctacgaatccaatacaaccgtgatccgactctctgcgcacgcttccaaatcaacatcggttc	131			
OY	72	ctacgaatccaatacaaccgtgatccgactctctgcgcacgcttccaaatcaacatcggttc	131			
Db	72	ctacgaatccaatacaaccgtgatccgactctctgcgcacgcttccaaatcaacatcggttc	131			
OY	132	taaaagtaactcgatccgatccgatacgaagaatcaatccagctgttcaatcgtgaatcttc	191			
Db	132	taaaagtaactcgatccgatccgatacgaagaatcaatccagctgttcaatcgtgaatcttc	191			
OY	192	caaaatcgaaagtatccctcgaaagaatgcatctgtatacaactatagttaagaactcttc	251			
Db	192	caaaatcgaaagtatccctcgaaagaatgcatctgtatacaactatagttaagaactcttc	251			
OY	252	caacctcctctcgtgaatccgtatcccgaaatctcaactccatcctctctgaacaatgtaata	311			
Db	252	caacctcctctcgtgaatccgtatcccgaaatctcaactccatcctctctgaacaatgtaata	311			
OY	312	caccatctcaactcgtcatggaagaacaattctgtttggaagatcatctctgaactaogtga	371			
Db	312	caccatctcaactcgtcatggaagaacaattctgtttggaagatcatctctgaactaogtga	371			
OY	372	aatcatctcggactcgcgcaggaacacccaggaatacaaacgcgcgtgtgtatcaaatc	431			
Db	372	aatcatctcggactcgcgcaggaacacccaggaatacaaacgcgcgtgtgtatcaaatc	431			
OY	432	tcaagatgatacaaatctctctgactacatacaaatccgcttggatcttcgttaccatacca	491			
Db	432	tcaagatgatacaaatctctctgactacatacaaatccgcttggatcttcgttaccatacca	491			
OY	492	tcgcttgataatactccaaatctacatacaagcgcgctcgtatccgcgcaggaacacgata	551			
Db	492	tcgcttgataatactccaaatctacatacaagcgcgctcgtatccgcgcaggaacacgata	551			
OY	552	caatctcggtaacaatccacgcgcttctaataacatcaatcgttcaaacctcgtgacgttgc	611			
Db	552	caatctcggtaacaatccacgcgcttctaataacatcaatcgttcaaacctcgtgacgttgc	611			
OY	612	cactcaccgcctacatctcgtatcaaatctcaatctcgttcgcgaacgaacctgaacgaata	671			
Db	612	cactcaccgcctacatctcgtatcaaatctcaatctcgttcgcgaacgaacctgaacgaata	671			
OY	672	agaaatccaagaacccgtgatacgaacaacccagtcacaaatctcgttatacctgaagac	731			
Db	672	agaaatccaagaacccgtgatacgaacaacccagtcacaaatctcgttatacctgaagac	731			
OY	732	tgaactaccctcgagtcgcaaaacccgttacaacttgcttgaaatctcgttcgatccgaacaa	791			
Db	732	tgaactaccctcgagtcgcaaaacccgttacaacttgcttgaaatctcgttcgatccgaacaa	791			
OY	792	cgltacgltcaacaatgtagatcgcgcgttctacatgtaacgtgaagaagtcgcgcgtgttc	851			
Db	792	cgltacgltcaacaatgtagatcgcgcgttctacatgtaacgtgaagaagtcgcgcgtgttc	851			
OY	852	tgttatgactacaaacatctaccctgaacatctctccctgtacgcgttgtaaccaaatcca	911			
Db	852	tgttatgactacaaacatctaccctgaacatctctccctgtacgcgttgtaaccaaatcca	911			
OY	912	caagaataacgcgcgtctgtttaacaagaacatacgttctgcaacaatgaatcgtgataagt	971			
Db	912	caagaataacgcgcgtctgtttaacaagaacatacgttctgcaacaatgaatcgtgataagt	971			
OY	972	caatgtctgagttaagaacaagaataacacgtcttggtctacaaatgcttcacagctgag	1031			
Db	972	caatgtctgagttaagaacaagaataacacgtcttggtctacaaatgcttcacagctgag	1031			
OY	1032	agaaaagatctctgctcgtcgtcgtgaatacccggaacgttgtaatactcgtccaggtag	1091			
Db	1032	agaaaagatctctgctcgtcgtcgtgaatacccggaacgttgtaatactcgtccaggtag	1091			

OY	1092	aatgaatacccaagaagccgacgggagtatactactaacaagaatgaatctgacggagaa	1151
Db	1092	AATGAATTCACCAAGAACGACACAGGATATCAATACAAATGCAAAATGAATCTGCAGACAA	1151
OY	1152	caatgtaacagatatcgggttcatctcggtttccaccaggttccaacatctcgtaaactggt	1211
Db	1152	CAATGTAACGATATCGGGTTTCATCTCGGTTTCACACAGTTCAACATATTCGTAACCTGGT	1211
OY	1212	tgcttccaactgtaacaaatcgtcagaatcgaaacgttccctctcgactctgggttgccttg	1271
Db	1212	TGCTTCCAACTGGTACAAATCGTCAATATGAAAGCTTCCTCGACTCTGGGTGCTCTTG	1271
OY	1272	ggagttcatcccggttgatgtaacggttgggaggaacgtctcggttaa	1317
Db	1272	GGAGTTCATCCGGTTGATGACGGTTGGGGTAACGATCGGCTGTAA	1317
RESULT	2		
LOCUS	AR169140	1330 bp	DNA
DEFINITION	Sequence 22 from patent US 6290960.		linear
ACCESSION	AR169140		
VERSION	AR169140.1	GI:17906909	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1330)		
AUTHORS	Kink, A., Thalley, B.S. and Stafford, D.C.		
TITLE	Vaccine and antitoxin for the treatment of C. difficile disease		
JOURNAL	Patent: US 6290960-A 22 18-SEP-2001;		
FEATURES	Location/Qualifiers		
source	1..1330		
BASE COUNT	400 a 339 c 246 g 345 t		
ORIGIN			
Query Match	98.6%;	Score 1304.4;	DB 6; Length 1330;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 1305; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
OY	12	gatgtcactccatcatgataatataagaacatacaatactcatctgaaacctgcg	71
Db	12	GGTGTCACCTTCACATGGAATACATCAAGAATCATATACCTCCATCCTGAACCTGGG	71
OY	72	ctaagaatccaatcaacctgtagctgacctgtctcgtctacgcttccaataatcaatcgtgtc	131
Db	72	CTACCAATTCCAATCACTCGATCGACCTGCTCTGCTACCGTTCCAAATCAACATCGGTTTC	131
OY	132	taaagttaactcgcgtatcgatcgacaagaagaatcagatccagctggttcaatctggaactcttc	191
Db	132	TAAAGTTAACTTCGATCCGATGACACAGATATAGATCCAGCTGTTCAATCTGGAATCTTC	191
OY	192	caaaatcgaaagtatactcctgaagaatgctatacgtlatatacaactatagtacgaanaactctc	251
Db	192	CAAAATCGAAGTTATCTCGAAGAATGCTATCTGTATACAACTGTATGACGAATACTCTTC	251
OY	252	caactcctcttgatctcgatataccgaaataacttcaactccatctctctgaaataatgata	311
Db	252	CACCTCTCTTCTGATCCGATATCCGAATATCTTCAACTCCATCTCTGMAACAATGAATA	311
OY	312	caacatatacaactgataggaaanaaatctcgtgttggaagaatatactctgaaactagctga	371
Db	312	CACCATATCATACATCGATGGAANAATTTCTGTTGGAAGATATCTCTGAACTACGGTGA	371
OY	372	aatactctggaactctgcaggaacatccaggaataccaacagcgtgtgtgatatccaatctac	431
Db	372	AATCATCTGGAATCTGCAAGGACATCTCAGGAATCAACAGCGCTGTTGTATTCAAATFAC	431
OY	432	tcaagatgatacaatctctgactatacaaatcgtcgtgatatctcgttataccaacaac	491

Db 432 TCAGATGATCAACATCTCTGACTACATCAATCCGCTGATCTGTTACATCAACCAACA 491
Qy 492 tctctgaataactccaaatactacatcaacggcgctctgaatcgacaagaacggtctc 551
Db 492 TCGTCTAATAAATCACTCAAAATCTACATCAACGCGCTGATGACACGAACCGATCTC 551
Qy 552 caatctggtaacatccacgctctctaataacatctgaatctgaacggtctgctga 611
Db 552 CAATCTGGTAACATCCACGCTCTCTAATACATCATCTTCAAACTGACGCTGTCTGTA 611
Qy 612 caactacgctacatctgatacaatactcaatctgtctgcacaagaactgaacga 671
Db 612 CACTCACCGCTACATCTGGATCAAAATCTTCAATCTGTTGACAAAGAACTGAACGAAA 671
Qy 672 agaatcaagaactctgaacgaacacgctcaatcttcgfatctcgtgaagaactctgg 731
Db 672 AGAAATCAAAAGACCTGTACACAAACAGTCCAAATCTGATCTGTAAGACTTCTGGG 731
Qy 732 tgaactactgcaatagacgaacacgctactacatgctgaatctgtaacgtaacaa 791
Db 732 TGACTACTGCACTAGACGAACAACCGTACTACTGCTGAATCTGATCTGACATCCGAACAATA 791
Qy 792 cgttgaacgtaacaatgtagtgcgggttacaatgtaacgtaacggtcgggttc 851
Db 792 CGTTGACGTCAACATGATAGGTATCCGCGTTACATGACTGTAAGGTCCGCGTGGTTC 851
Qy 852 tttatctacccaacatctacactgaactcttcctctgacggtgtaacgaatctcat 911
Db 852 TGTATATCTACCAACATCTCACTGAACTCTCCCTGATCCGTTGATGCAAAATTCATCAT 911
Qy 912 caagaaatacgcgtctgtaacgaagaacatctgctgcacaacatgatactgatacat 971
Db 912 CAAGAATACGCGCTGTGTAAACAAGCAATATCGTTGCGAACAATGATGTGATATCAT 971
Qy 972 caatgtgtatgaagaacgaatacgcgtctggtaccatgctctcaaggctgtgt 1031
Db 972 CAATGTGTATGTTAAGAAACAAGAAATACCGTCTGGCTACCAATGCTTCTCAGCGCTGTT 1031
Qy 1032 agaaaagatcttctctctctgtaaatcccggaacggttgtaactctgctgaagtgt 1091
Db 1032 AGAAAAGATCTGTCTCTCTGGAATCCCGGAGGTGTGTAATCTGCTCAAGGATGTGT 1091
Qy 1092 aatgaataccaagaacgacaggtatcactaaacaaatgcaaaatgatactctgcagaa 1151
Db 1092 AATGAATCCAAAGAACGACGAGGTATCATAACAATGCAAAATGATCTGACGAGCAA 1151
Qy 1152 caatgtgaacgatacgggttcatcgggttcacacagttcaacaaatacgttaactgt 1211
Db 1152 CAATGTGATGATATCGGTTTCATCGGTTTCCACACAGTTCAACAAATATCGCTAAACTGTT 1211
Qy 1212 tgcctcaactgtgataatctgacatcgaaactctctctcgaactctggtgtgtcttg 1271
Db 1212 TGCTTCCAACTGTGATCATGTCTGATCGAAGCGTTCTCTGCGACCTGTGGGTGCTCTTG 1271
Qy 1272 ggaattcatccggttgatgaacggttgagggtgaacgctcgtgtaa 1317
Db 1272 GGAGTTCAATCCGCTGATGACGTTGGGTTGAACGCTCCGCTGTAA 1317

RESULT 3
LOCUS AX036243 1330 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 22 from Patent EP1041149.
ACCESSION AX036243
VERSION AX036243.1 GI:11225862
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 1330)
AUTHORS Kink,J.A., Flicha,D.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and
TITLE Vaccine and antitoxin for treatment and prevention of C. Difficile

JOURNAL disease
Patent: EP 1041149-A 22 04-OCT-2000;
OPHIDIAN PHARM INC (US)
FEATURES
source Location/Qualifiers
1..1330
/organism="synthetic construct"
/db_xref="taxon:32630"
1..1317
/note="Synthetic"
/codon_start=1
/transl_table=11
/protein_id="CAC16478.1"
/db_xref="GI:11225863"
CDS
translation="MARLSTFEYIKNIINTSIILNRYESNHLIDLSRVASKINIGS
KYNPDIDKNOIOLNLESSKLEVLKNAIYVNSMVENPSTSFRIKPYNSISLN
EYTIINCMENNSGKRVSLAYGEIITLDTDTGRIKRVYFKYSQMINIDYINRAIYFT
ITNNRLNRSKITIINGLIDQKFIISNGNTHASNIMFPLDGRDTHRTIWKTFELFD
KELNEKEIKDLYDNOSNSGILKDFWGDYLYDXRPYMLNLDPNKYVYVNNVIGYIM
YLKGRGSVMTNINILNSLYRGTFEIIKRYASGKNDRVNRNDVYINVVKNKEFR
LATNASQAVEKILSALIEIPDVGNLSQVVMKSKMDGITTNCCKNLDNDNGNDIGFI
GFHOENIATKLIVASNMVNRQIRRSSTLTGCSNEFIIPVDGMEERPL"

BASE COUNT 400 a 339 c 246 g 345 t
ORIGIN

Query Match 98.6%; Score 1304.4; DB 6; Length 1330;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 gatgtctaccttactgaatacatcaagaacatcactccatccctgaactgcg 71
Db 12 GCTGTACTCTCAGTAATACATCAAGAATCATCATCAATCTCCATCTCACTGCGG 71
Qy 72 ctacgaatccaatccactgatacgtctgctgctacgtctccaaatacaactcgttc 131
Db 72 CTACGAATCCAAATCACCCTGATTCGACCTGTCTGCTACGCTTCCAAATCAACATCGGTTTC 131
Qy 132 taaagttaacttgcattccgctgcagaagaatcagatccacgctgttcaatctgaactctc 191
Db 132 TAAAGTTAACTTGGATCCGATCGATCAACAAGATTCAGACTTTTAATCTGGAATCTTC 191
Qy 132 TAAAGTTAACTTGGATCCGATCGATCAACAAGATTCAGACTTTTAATCTGGAATCTTC 191
Db 132 TAAAGTTAACTTGGATCCGATCGATCAACAAGATTCAGACTTTTAATCTGGAATCTTC 191
Qy 192 caaatcgaagtataccttgaagaatgctatcgtatacaactctatgtaacgaactctc 251
Db 192 CAAAATGAGATTATCTCTGAAGAATGCTATGATCAACTCTATGACGAAGAATCTCTC 251
Qy 252 caactcctcttgatccgtatcccgaaatacttcaactcactctctctgaacaaatga 311
Db 252 CACCTCTTCTGATCCGATCCCGAATATCTCAACTCCATCTCTCTGAAACAATGATA 311
Qy 312 caccatcatcaactgacatggaagaacatctgttggaagaatctctctgaactcgtga 371
Db 312 CACCATCATCAACTGCATGGAAGAAACAATCTGTTGGAAGATATCTCTGACATCACTGTA 371
Qy 372 aatcatctgactctgcagaacacatcagaacatcaaacagcgtgtgtatcacaactac 431
Db 372 AATCATCTGACACTCTGAGGACACTCAGGAATCAACAACGCTGTGATTAATAATATC 431
Qy 432 tcaagatgaatacaatctgactaactcaatctgctgtgattcttgattacatcaacaa 491
Db 432 TCAGATGATCAACATCTCTGATCAATCAATCCCTGATCTGTTACCATCAACAA 491
Qy 492 tctctgaataactccaaatactacatcaacggtcgtctgaatcgacaagaacggtctc 551
Db 492 TCGTCTAATAAATCACTCAAAATCTACATCAACGCGCTGATGACACGAACCGATCTC 551
Qy 552 caatctggtaacatccacgctctctaataacatctgaatctgaacggtctgctga 611
Db 552 CAATCTGGTAACATCCACGCTCTCTAATACATCATCTTCAAACTGACGCTGTCTGTA 611
Qy 612 caactacgctacatctgatacaatactcaatctgtctgcacaagaactgaacga 671
Db 612 CACTCACCGCTACATCTGGATCAAAATCTTCAATCTGTTGACAAAGAACTGAACGAAA 671

QY 672 agaaatcaagaacctgtacgacaacacagtcaccaatctgtgtatccttgaaagacctctggg 731
 Db 672 AGAATATCAAAAGCCTGTACGACAACCAAGTCCATTTGGTATCCCTGAAAGACTTCTGGGG 731
 QY 732 tgactactcagctacgacaacacagctactactatctgaatctgtacgactcgcgaacaata 791
 Db 732 TGACTACTCGAGTACGACAAACCGTACTACATGCTGATGTGACGATCCGAAACAATA 791
 QY 792 cgttgacgtcaacaatgtaggtatccggtttacatgtacttgaaaggtccgctgtgtc 851
 Db 792 CGTGTAGCTCAACATGTAGGTATCCGGGTTACATGTACTGAAAGTCCGGCTGGTGT 851
 QY 852 tcttatgactacccaacatctactcgtgaactcttccctgtacccgtgtgtacccaatctcat 911
 Db 852 TGTATGTACTACCAACATCTACTCTGAACTCTCCCTGTACCGGTGTACCAATTCATCAT 911
 QY 912 caagaataacgctctgtgttaacagaagacaataatgcttcgacaagaatgtatgtatcat 971
 Db 912 CAGCAATACCGCTGTGTGTACAAAGGACAATATGCTTGCAACAATGATCGTATACAT 971
 QY 972 caatgtgtgtatgaagaacaagaataacgctctgtgtacccaatgtcttcacagctgtgt 1031
 Db 972 CAATGTGTAGTTAAGAACAAAGATACCGTCTGGCTACCAATGCTTCTCAGGCTGGTGT 1031
 QY 1032 agaaaagatctgtctgtctcgtcgtgaatcccggaacgtgtgtatctgtctcaagtaagt 1091
 Db 1032 AGAAAAGATCTGTCTGTCTGTGGAATCCCGAGCTTGGTATGTTCTCAGTGTAGTGT 1091
 QY 1092 aatgaataacgaagaacgacaggggtatcacttaacaataatgaatgaatcgtcgaagaca 1151
 Db 1092 AATGAATAACGAAGACGACGAGGTATCTACTACAAATGCAAAATGATCTGCGAGACAA 1151
 QY 1152 caatgtgaacatctcgtgttccatcgtgttccaccagttccacaataatcgtcgaactgt 1211
 Db 1152 CAATGTGAACATCTCGTGTTCATCGGTTTCACACAGTTCAACAATATCGCTAAACTGT 1211
 QY 1212 tgcttccacagctgtacaaatcgtcagatcgaaacgttccctcgtcgaactcgtgtgtct 1271
 Db 1212 TGCTTCCACAGTGTACAAATCTGTGGAATCCCGAGCTTGTCTCAGTGTAGTGT 1271
 QY 1272 ggaatctacccggttgcgtacggttgggggtgaacgctcgcgtctga 1317
 Db 1272 GGAGTCACTCCGGTTGATGACGGTGGGGTGAACGTCCTCGCTGTGA 1317
 ESURT 4
 LOCUS XU22962 1338 bp DNA linear SYN 08-NOV-1995
 DEFINITION Synthetic botulinum neurotoxin serotype A Hc fragment (botA) gene,
 complete cds.
 ACCESSION U22962.1 GI:733428
 VERSION U22962.1
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (bases 1 to 1338)
 AUTHORS Clayton, M.A., Clayton, J.M., Brown, D.R. and Middlebrook, J.L.
 TITLE Protective vaccination with a recombinant fragment of Clostridium
 botulinum neurotoxin serotype A expressed from a synthetic gene in
 Escherichia coli
 JOURNAL Infect. Immun. 63 (7), 2738-2742 (1995)
 MEDLINE 95310035
 REFERENCE 2 (bases 1 to 1338)
 AUTHORS Brown, D.R.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-1995) Douglas R. Brown, Toxinology, U. S. Army
 Medical Research Institute of Infectious Diseases, Bldg. 1425, Fort
 Detrick, Frederick, MD 21702-5011, USA
 FEATURES
 source 1..1338
 /organism="synthetic construct"
 /db_xref="taxon:32650"

/lab_host="Escherichia coli"
 /note="based on Clostridium botulinum Type A neurotoxin
 sequence"
 gene 9..1325
 /gene="botA"
 CDS 9..1325
 /gene="botA"
 /codon_start=1
 /transl_table=11
 /product="botulinum neurotoxin serotype A Hc fragment"
 /protein_id="AA80610.1"
 /db_xref="GI:733428"
 /translation="MARLSTPEYIKNIINISLNRRESNHLIDLSRYAKINIGS
 KVNEDIDKNQIOLEFMLESSKLEVLINAIIVNSMTEPSTSFWIRIPKYSISLN
 EYIINCMENNSGMKVSILNYGELIITLQDTEIKQRYVKYSQMINISDYINRMIVT
 ITNNRNSKIYINGRLIDQKPLISGNTHASNMIFKLDGCRDRIYIWKYFLFD
 KELNEKEIDIVDNOSNGILDFMDYLOQPYMLLYDPNRYVDVNNYNGIGYM
 YLKPFGSVHTTIYIYINSSLYRTRKTIIRKYSAGKNDIVRNDRYIVNYYKNEXR
 LATNAGQAEVKTLNLEIPDGNLSQVYVMSKNDQITNCKMHLQDNNQNDIGFI
 GFHFNNAIKLVASNMNRYNRIEKSSTFLCSWEFLPVDGWERPL"
 BASE COUNT 401 a 342 c 249 g 346 t
 ORIGIN
 Query Match 98.6%; Score 1304.4; DB 12; Length 1338;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1305; Conservatly 0; Mismatches 1; Indels 0; Gaps 0;
 QY 12 gatgtacaccttcaactaataacgaagaatcatatcatcactcctcgaacctg 71
 Db 20 GCTGTCTACTTCTCCTAATATACATCAAGAACATCATATACCTCCCTGTGAACCTGGC 79
 QY 72 ctacgaatccaatcacctcgtacgtgtcgtcgtacgtgtcgaataacacacacgtgtc 131
 Db 80 CTACGAATCCAATCACCTGTATCGACGTCTCGCTACGTTCCAAATCAACATCGGTTG 139
 QY 132 taaagttaactcgtatccgtacgcgaagaatcgaatcgaatcgtgttcaatcgtgaacttc 191
 Db 140 TAAAGTTAACTTCGATCCGATCGATCGACAAAGAAATCGATCCAGCTGTCAATCTGAAATCTTC 199
 QY 192 caaatcgaagtatccctcgaagaatgtctatcgtatatacaactctatgtacgaagaactctc 251
 Db 200 CAAATATGAAGTATTCCTGGAAGATGCTATCGTATACACTATATAGCAAAACTTCTC 259
 QY 252 caactccttctgtatccgtatccgaaatacttcaactcactcctcctcgaagaatgaata 311
 Db 260 CACCTCTTCTGTGATCCGATCCGGAATCTTCAACTCATCTCTGTGAACATGATA 319
 QY 312 caccatatacaactcgtacgtgaagaacaatctgtgtggaagtatcctcgtacgaactgtga 371
 Db 320 CACCATATCAACTGCATGGAAACAAATCTGTGGTGAAGATCTCTGAACTACGCTGA 379
 QY 372 aatcatctgacctcgtcaggaacactcaggaatcaacaacagcgtgtgtatcacaatactc 431
 Db 380 AATCATCTGGAGCTGTGAGGACACTCGAAGAAATCAACACGAGTGTATTCAAATATCTC 439
 QY 432 tcagaatgaacaactcctcgtactacatcaatcgtcgtatcctcgttaccatcacaaca 491
 Db 440 TCAATGATGCAACATCTGTGATCAATCAATCCCTGATCTGTGTTACATCAACAA 499
 QY 492 tctgtctgaataactcgaataatcacaatcgaacgacctgtatcgtacgaagaacagatctc 551
 Db 500 TCGTCTGAATTAACCTCCAAATCTACATCAACGCGCTGTATGACCAAGAACGATCTC 559
 QY 552 caatcgtgtgaacatcgaacctctaaataacatcatgttcaaaactggaacgtgtgtgtga 611
 Db 560 CAATCTGGGTAACTCCACCTCTTATTAACATCATATTTAAACTGGACAGCTTGTCTGA 619
 QY 612 cactcaccgtctacatcgtgaatcaataactcctcgttctcgaagaagaactgtgaaga 671
 Db 620 CACTCACCGCTACATCTGATCAAAATCTCTCATCTGTGTGACAAAGAACTGAACGAA 679
 QY 672 agaaatcaagaacctgtacgacaacacagtcaccaatctcgtatccttgaaagacctctggg 731

|||||
Db 680 AGAATCAAGACCTGTACGACACACAGTCCAAATCTGTCTCTGAAAGACTTGGGG 739
OY 732 tgaactacgtacgtacgaacaaacgtactacatctgaatctgtacatctcgaaacata 791
Db 740 TGACTACCTGAGTACGACCAACCGTACTACATCTGTAATCTGTACGATCCGAACAATA 799
OY 792 cgttgaacgtacaacatgtaagtlaccgcggttatactgatacgtgaagtcgcggtgtc 851
Db 800 CGTTGACGTCAACATGTAGTATCCGCGTTACATGTACTGTAAGGTCCGCGTGGTTC 859
OY 852 tgttgaactacaacatctactcctgaactcttcctgtlaccgtygtlaccacaatcatcat 911
Db 860 TGTATACTACCAACATCTACTGTAAGTCTTCCCTGATACCGTGTACCAAAATTCATCAT 919
OY 912 caaagaatacgcgtctgttgaacaaagaaataatcgttccgaaacaaatgatacgtatatacat 971
Db 920 CAAGAAATACGCGTGTGTGTAACAAGCAATATGCTTCGCAAAATGATCGTGTATACAT 979
OY 972 caatgttgaagtlagaacaaagaataccggtctgtgtaccacatgtcttcaggctgtgt 1031
Db 980 CAATGTTGTACTTAAGAACAAAGATACGCTGTGCTACCAATCTTCTCAGCGTGGTGT 1039
OY 1032 agaaaagatctgtctgtctgtgaatcccgagctgtgtatctgtctcagtagtgt 1091
Db 1040 AGAAAGATCTTGTCTGTCTGTGAATCCCGAGCTTGTGTATCTGTCTGAGTAGTGT 1099
OY 1092 aatgaatccaaagaacgaacgggtatcaatacaaatgcaaaatgaaatctgcagaaca 1151
Db 1100 AATGAATCCAAAGAACGACGAGGTATCTACATACAAATGCAATGATCTGCGAGGACAA 1159
OY 1152 caatgttgaacgatacgtttcaccggtttccaccagttccacaatactgcataacgtgt 1211
Db 1160 CAATGTGTACATATCGGTTTCATCGGTTTCACACAGTTCAACATATTCGTTAAACGTGT 1219
OY 1212 tgccttcaactgtgtacaaatcgttcagatcgaacggttccctcgtcagctgtgtgtcttg 1271
Db 1220 TGTCTCAACGTGTACATCGTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1279
OY 1272 ggaattatcccggttgaatgaacggttgggtgaacggtccgtgttaa 1317
Db 1280 GGAATTCATCCCGGTGTGATGACGGTGTGGGTGAACGTCCTGCTGA 1325

RESULT 5
AR000030 1402 bp DNA linear PAT 04-DEC-1998
LOCUS AR000030
DEFINITION Sequence 25 from patent US 5736139.
ACCESSION AR000030
VERSION AR000030.1 GI:3962561
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.
TITLE Treatment of Clostridium difficile induced disease
JOURNAL Patent: US 5736139-A 25 07-APR-1998.
FEATURES
Location/Qualifiers
source
1..1402
BASE COUNT 420 a 360 c 260 g 362 t
ORIGIN

Query Match 98.6%; Score 1304.4; DB 6; Length 1402;
Best Local Similarity 99.9%; Pired. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 12 gatcttacctgaactgaataatcaagaacataataatccatccatccgtgaacgtgcg 71
Db 84 GCTGTACCTTCACTGATATCATCAAGAACATCATCATCAATCTCTGAAACGTCGG 143

OY 72 ctacgaatccaatcaactgtatcgacgtgtctcgtaacgttccaaatcaacatcggttc 131
Db 144 CTAGGATTCATACCTGATGACCGTGTCTGCTACGCTTCCAAATATCAACATCGGTTTC 203
OY 132 taaagttaactcgatcgatcgacaagaatcagatccagctgttcaatctgtaacttc 191
Db 204 TAAAGTTAATCTGATCCGATCGACAAAGATGAGATCCAGCTGTTCATCTGGAATCTTC 263
OY 192 caaatcgaaatctatccgtgaagaatgctatcgatcaactctatgtaagaactcttc 251
Db 264 CAATATGAAAGTTATCTTGAGAAATGCTATGATACAACTTATGTATGACAAATCTTCTC 323
OY 252 caccctctctgtatccgtatcccgaaatactccaactccatctctctgacaataa 311
Db 324 CACCTCTTGTGATTCGGTATCCCGAAATCTTCACTCCATCTCTGTGACATGAAAT 383
OY 312 caccatcaactcgaatgtgaaacaaatctgtgtgaaagtatctctgaaatcggtga 371
Db 384 CACCATCATCACTGATGGAACAAATCTGCTGGAAGTATCTGTAATCTGTAATCGGTGA 443
OY 372 aatcatctgacgtctgagaacaaatcaagaacaaacagcgtgtgtatcaaatctc 431
Db 444 AATCATCTGGACCTTGTGACGACACATCAGAAATCAACACGCTGTGTATCAAAATCTC 503
OY 432 tcaatgatcaacatctctgaactacaatcaatcgtatgtctgttaccatacacaaca 491
Db 504 TCAGATATATCAATCTCTGATCAATCAATCGCTGATCTTGTGTTACCAATCCACACAA 563
OY 492 tctgtgaatcaactccaanaatctacatcaacagcgtctgtatcgaccagaacagatctc 551
Db 564 TCGTGTGAATATCTCCAAATCTATCATCAACGCGCTGTGATGACACAGAAACGATCTC 623
OY 552 caatctgtgtgaatccacacgtcttcaataacatcaatglttcaactgtgacgtgtgtga 611
Db 624 CAATCTGGGTAAATCCACGCTTCTTAATATCATCATGTTTCAAACTGACGCTGTGTGA 683
OY 612 caatcaccgtatcatctgatacaataactcaatctgttctgcaagaagaactgtgaacaa 671
Db 684 CACTCACGCTTACATCTGATCGATCAATATCTTCAATCTGTTGACAAAGAACTGACGAAA 743
OY 672 agaaatcaagaacctgtgaacaaacagtcgaatctgtgtatctcgtgaagaactctgggg 731
Db 744 AGAAATCAAGACCTGTACGACCAACAGTCCAAATCTGATGATCTGAAAGACTCTTGGGG 803
OY 732 tgaactacgtacgtacgaacaaacgtactacatgctgaatctgttaacatccgaacaaata 791
Db 804 TGACTACCTGACGATGACGACAAACCGTACTGATGCTGAATCTGTACGATCCGAACAAATA 863
OY 792 cgttgaacgtacaacatgtaagtlaccgcggttatactgatacgtgaagtcgcggtgtc 851
Db 864 CGTTGACGTCAACATGTAGTATCCGCGTTTACATGTACTGTAAGGTCCGCGTGGTTC 923
OY 852 tgttgaactacaacatctactcctgaactcttcctgtlaccgtygtlaccacaatcatcat 911
Db 924 TGTATGACTACCAACATCTACTGTAAGTCTTCCCTGATACCGGTGACCAAAATTCATCAT 983
OY 912 caaagaatacgcgtctgttgaacaaagaataatcgttccgaaacaaatgatacgtatatacat 971
Db 984 CAAGAAATACGCGTGTGTGTAACAAGCAATATGCTTGTGCAAAATGATGCTGTATACAT 1043
OY 972 caatgttgaagtlagaacaaagaataacggtctgtgtaccacatgtcttcaggctgtgt 1031
Db 1044 CAATGTTGTGTATGATTAAGAACAAAGATATGCTGTGCTACCAATGCTTTCAGCGTGGTGT 1103
OY 1032 agaaaagatctgtctgtctgtgaatcccgagcgttgtgtaactgttctcagtagtgt 1091
Db 1104 AGAAAGATCTTGTCTGTCTGTGAAATCCCGAGCTGTGTAATCTGTCTCAGGTAATGTTGT 1163
OY 1092 aatgaatccaaagaacgaacgggtatcaatacaaatgcaaaatgaaatctgtgaagcaaa 1151
Db 1164 AATGAATCCAAAGAACGACGAGGTATCTAATCAAAATGCAAAATGCAAAATGCAAGACAA 1223
OY 1152 caatgttgaacgatacgtttcaccaggttccaccagttcaacaatacgtctaaactgt 1211

Db 1224 CAATGCTAACGATATCGGTTTCATCGGTTTCACACAGTTCAACAATATCGCTAAACGTGT 1283
 Oy 1212 tgcctccacacgtgtaacacgttcagatcgacagcttcctcctcgcactcgtggttcctctg 1271
 Db 1284 TGCTTCCACACTGATACATCTCTCAGATCGACAGTTCCTCTGCACTCGGGTTCCTCTTG 1343
 Oy 1272 ggaattcaaccgggttgatgacgggttggtgtaacgctcgcgtgtaa 1317
 Db 1344 GGAGTTCATCCCGGTTGATGACGGTTGGGGTGAAAGTCCTCGGTAA 1389

RESULT 6
 ARI69141
 LOCUS ARI69141 1402 bp DNA linear PAT 17-DEC-2001
 DEFINITION Sequence 25 from patent US 6290960.
 ACCESSION ARI69141
 VERSION ARI69141.1 GI:17906911
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1402)
 AUTHORS Kink,J.A., Thalley,B.S. and Stafford,D.C.
 TITLE Vaccine and antitoxin for the treatment of C. difficile disease
 JOURNAL Patent: US 6290960-A 25 18-SEP-2001;
 FEATURES location/Qualifiers
 source 1..1402
 BASE COUNT 420 a 360 c 260 g 362 t
 ORIGIN

Query Match 98.6%; Score 1304.4; DB 6; Length 1402;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 12 gatgtctactccttcacgtatcatcaagaacatcatcaatacctccatcctcgtgaacctg 71
 Db 84 GCTGTCTACTCTCACTGATACATCAAGAACATCATCATCAATCTCATCTCGAACCTCG 143
 Oy 72 ctagaatccaatacctacgtacgtacgtctcgtctgaacctcccaaaatcaacatcgtgtc 131
 Db 144 CTAGAAATCCATACCTGATCGATCGACCTGTCTGCTACCTTCCAAAATCAACATCGGTT 203
 Oy 132 taaagttaactcgatccgatccgacgaagaatcagatccagctgttcaactcgtgaactctc 191
 Db 204 TAAAGTTAACTTCGATCGATCGACAGAAATCAGATCCAGCTGTTCATCTGGAATCTTC 263
 Oy 192 caaaatcgaaagtatcctgaagaatgctatcgtatcaactctatgttgaagaactctc 251
 Db 264 CAAATCGAAGTTATCCGAGAAATGCTATGATACAACTATGATGCAAAACTCTCTC 323
 Oy 252 cactcctctcgtgacgtatccgccgaataactcaactccatcctctcgtgaacaatgtaata 311
 Db 324 CACTCTCTTCTGAGATCCGATATCCGAAATACTTCAACTCCATCTCTGGAACATGATA 383
 Oy 312 caccatcaacaacgtacatggaagaactcgtgttggaagtatcctcgtgaactcgtgta 371
 Db 384 CACATATCATCACTGATCGATCGACAGAAATCTTGTTGGAATATCTCTGGAATCGGGA 443
 Oy 372 aatcatctgagctctgcaagacactcaggaatacaaacagcgtgttlatcaaatc 431
 Db 444 AATATCTGGAGCTCTGCGAGCACTCAAGGAATCAAAAGCGTGTGATTCATTAATCTC 503
 Oy 432 tcaatgatatacaactctcgtactcaatcactcgtcgtgactcgttaccatcacacaata 491
 Db 504 TCAAGATGATCAACATCTCTGATCAATCAATCGCTGATCTCTGTTACCATCCACCAAAA 563
 Oy 492 tgcgtctgataactccaaaactcatcatcaagcgcgtctgactgacccaagaacgatc 551
 Db 564 TCGTCTGATTAATCCAAAATCTATCAACGCGCTGTGTGATCGAACGAACCGATCTC 623

Oy 552 caatctggtgaacatccacgcttctaaatcatcatgttcaaatcgtgacggttgcgtgta 611
 Db 624 CAATCTGGATTAATCAACCGCTTCTAATATCATCTGTCAAACTGGACGTTGCTGA 683
 Oy 612 cactcaccgcttaactcctggtatcaataactcaatcgttctgacaagaactcgaagaaa 671
 Db 684 CACTCACCGCTTACATCTGAGATCAATCTTCAATCTGTTGACAAAGAACTGAAACGAAA 743
 Oy 672 agaatcaagaacgttgaacagaacacagtcgaattcctgcatcctgaaagactctggg 731
 Db 744 AGAATATCAAGACCTGTATGACACAAACGATCAATCTGATCTCTGAAACACTCTGGGG 803
 Oy 732 tgactactcgtcagtaagacaacacgctactcaatgctgaatcgttgaatcgcgaacaata 791
 Db 804 TGACTTACTCTGACATGACACAAACCGTACTCATGCTGAATCTGTACATCTCGAACAATA 863
 Oy 792 cgttgaagctaaacatgttaagttacgcggttcaatgactcctaagaagctcgggtgtc 851
 Db 864 CGTTGACGTCACAAATATAGGTATCCCGGTTTACATATCTGACAAAGTCCGGTGGTTC 923
 Oy 852 tgtatgactacaacaactcaactgaactcttccctgtacacgttgcatacaatcatcat 911
 Db 924 TGTTATGACTACCAACATCTAATCTGAACTCTTCCCTGTACCGTGTACCAAAATTCATCAT 983
 Oy 912 caagaatacgcgctcgtgtgaacaagaacatlatcgtcgaacaatgacgtgtatacat 971
 Db 984 CAAGAATATACGCGCTCTGTGAACAAGCAATATGTTGCGCAACAAATGATGTTATCAT 1043
 Oy 972 caatgtgtgattgaagaacaagaatacgcgttcggtcaccgaatgctcctaagctgtgt 1031
 Db 1044 CAATGTTGATTGATGAACAAAGATATCCGCTGTGGCTTACCAATGCTTCTCAAGCTGTGT 1103
 Oy 1032 agaaagaactctgtcgtcgtcgtgaatcccggaacgttgcgtgaatcgtcctaagctgtgt 1091
 Db 1104 AGAAAGATCTGTTCTGCTCTGTGGAATCCCGGAGGTGTGAATCTGCTCAGAGTGTGT 1163
 Oy 1092 aatgaatccaagaacgacgaaggtatcactaacaatgcaaatgtaattcgaagacaa 1151
 Db 1164 AATGAATCCAAAGAACACCAAGGATATCACTAACAAATGCAAAATGTGACAGGCAA 1223
 Oy 1152 caatgtgaagatatacgtttcatcgtttccaccagaattcaacaatacgtcctaactgtgt 1211
 Db 1224 CAATGTTAGATATCGGTTTCATCGGTTTCCACCACTTAAACATATCTCTAACTGCT 1283
 Oy 1212 tgcctccacacgtgtaacacgttcagatccgacgttccctcctcgcactcgtggttgcctctg 1271
 Db 1284 TGCTTCCAACTGATACATCTCTCAGATCGACAGTTCCTCTGCACTCGGGTTCCTCTTG 1343
 Oy 1272 ggaattcaaccgggttgatgacgggttggtgtaacgctcgcgtgtaa 1317
 Db 1344 GGAGTTCATCCCGGTTGATGACGGTTGGGGTGAAAGTCCTCGGTAA 1389

RESULT 7
 AX036246
 LOCUS AX036246 1402 bp DNA linear PAT 16-NOV-2000
 DEFINITION Sequence 25 from Patent EP1041149.
 ACCESSION AX036246
 VERSION AX036246.1 GI:11225864
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 1402)
 AUTHORS Kink,J.A., Firca,J.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and Williams,J.A.
 TITLE Vaccine and antitoxin for treatment and prevention of C. Difficile disease
 JOURNAL Patent: EP 1041149-A 25 04-OCT-2000;
 FEATURES OPHEIDIAN PHARM INC (US)
 source location/Qualifiers
 1..1402
 /organism="synthetic construct"

```

CDS
1..1389
/db_xref="taxon:32630"
/feature="Synthetic"
/codon_start=1
/transl_table=11
/protein_id="CAC16479.1"
/db_xref="GI:11225865"
/translation="MGHHHHHHHHSSGHIEGRHMASMARLLSTFTEYTIKNIINSTI
LNLLENSENHLIDLSRYASKINIGSKVNPFDLPDKQIOLFINESSKLEIVILKALNAVYMS
MEYNESTFWIRIKRYPNISILNNEYIINCMEHNSGKMSYLANGELITWLPDQCEIKR
ORVYKFSOMINISDIYNRMLEFVITNNRLNKSITNGRILDDKPTSLNIGNHASNN
IMFLDGCSDPTHRITWIKFENLFDKELNEKIKOLYNQNSGILKDFMGVIOYDKR
YYMNLIDPKKYVDVNVNNGVIRGMYLFGPGRGSVWNTIYNSSITLPGCTFLIKYVAGS
KNKDIIVNNDRVYINYYKKREYRLATNAGSAGEYKILSALEIYDVGNLSOYYVMKSK
NDOSITNCKKCNLDQNNQNDIGIFIGFQFQFNIAKLAVASNMNIRDIERSRILGCSWEF
IPVDGIGPERPL"
BASE COUNT      420 a      360 c      260 g      362 t
ORIGIN

```

Query Match	98.6%	Score 1304.4	DB 6	Length 1402
Best Local Similarity	99.9%	Prod. No. 0		
Matches 1305	Conservative	0	Mismatches 1	Indels 0
				Gaps 0
QY 12	gatgtctaccttcacatgatatacatcaagaacatacctaatactccatccctgaacctgcg	71		
Db 84	GCTGCTACCTTCACTGAAATACATAAAGACATCTCAATACCTTCATCTCGAACCTTCGG	143		
QY 72	ctacgaatccaatccaacctgatcgacctgctctgcgtacgcttcccaaatccaatcgtgttc	131		
Db 144	CTACGAATCCCAATCACTGATCGACCTGCTCTCGTACGCTTCCAAATCAACATCGGTTC	203		
QY 132	taaaagttaactcgatccgcgatcgaaagaatcagaatccaagctggttccaatcgtgaatcttc	191		
Db 204	TAAAGTTAACTTCCTGATCCGATCGAACAAATAATCAGATCCAGCGTTCATCTGGAAATCTTC	263		
QY 192	caaaatcgaaagtataccctggaagaatcgcatacgtatacaacctatgatcgaaacctcttc	251		
Db 264	CAAAATCGAAGTTATCCCTGGAAGAAATGCTATCTGATACAACTCTATGTACGAAACTTCTC	323		
QY 252	caacctctcttcgtagatccglatcccgaaatctaactccaatcctctctgaaacaatgtaata	311		
Db 324	CACCCCTCTTGGATGCCGTATCCGAATTACTTCAATCCATCTCTGTAACAAATGATA	383		
QY 312	caacgatccaacttcataaggaaaacaaatcttggtttggaagaatctctcgaactaagtgga	371		
Db 384	CACCACTCATCAACTGCATGAGAAACAAATCTGGTTGGTGAAGATCTTGGAATCTAGGTGA	443		
QY 372	aatactctcgagactcgcaggaacactccaggaatacacaagaacgctgtgttatccaatctc	431		
Db 444	AATCATCTGAGACTCTGCAGAGACACTCAGGAATATCAACAGCGCTGTGTATTCATAATCTCTC	503		
QY 432	tcaagatgatacaacatctctgtaactacataatcgcgtgatatcttgttaccatccaacaa	491		
Db 504	TCAAGTATGATCAACATCTCTGACTACATCAATCCGCTGGATCTTCTGTACCATCAACCAAA	563		
QY 492	tcgcttgataaactccaacatctacataaagcgccgcttgatcggccggaaacgaatctc	551		
Db 564	TGCTGTGATTAATCTCCAAATATTAATATTAACGCGCGCTGTATGATCGACCAAGAACCCATCTC	623		
QY 552	caatctgggtlaacataccaacgcttccaataacaatacgttccaactcgaacgagctgtgcgtga	611		
Db 624	CAATCTGGGTAAACATCCACACGGCTTAAATTAACATCATGTTCAAACAGCGAAGCGTGTCTGTA	683		
QY 612	caactacacgctacatctcgtgatccaatctaactccaatctgttcgacaaagaactgaacgaaaa	671		
Db 684	CACATCACGCTTAACATCTGGATCAATAATTAATCTTCATCTGTTCGCAAAAGAACATGAAAGAAA	743		
QY 672	agaaatacaagaacccgtgtacgaacacacagttccaatctcgttatccctggaagaagcttttggg	731		
Db 744	AGAAATCAAAAGCCCTGTATCGACACACCAAGTCCATTTCTGTGATCTCGAAAGAACTTTCTGGGG	803		
QY 732	tgaactactgcagtaacgaacaaacgcttactacatgcttgatatctgtacgataccgaacata	791		

D	b	804	TGACTACCTGCAGTATCGCAACAACCGTACTACATGTGTGAATCTGTACGATTCGGAAACAATA	863
Q	y	792	cgttgaagctcaacaatgtagtgatccgcgggttaacatgtaactgtgaaggctcggatggttc	851
D	b	864	CATTGACGCTCAACAATGATGAGTATCCGCCGTTTACATGTACCTGGAAGAGGTCCGCGTGTC	923
Q	y	852	tgttatgctaccacaatactaccctgaactcttcctctgacctgtaacggttgtlaccaatlcatcat	911
D	b	924	TGTTATGACTCACCACAATCTACCTGAACCTCTCCCTGTACCGTGGTACCAAAATTCATCAT	983
Q	y	912	caagaataacgcgctctgtttaacaaggaacatctcgctcgacaacatgactgtgatatacat	971
D	b	984	CAGCAAAATACCGCGTCTGGTAACAGGACAATATCGTTCCGACAACATGATCGTGTACAT	1043
Q	y	972	caatgtttgtagttaagaacaagaataaccgctctggctgtaaccaatgctltccagctgtgt	1031
D	b	1044	CAATGTGTAGTTAAGAACAAGAANTATACCGTGGCTACCAATGGCTTCTCAGGCGTGCGT	1103
Q	y	1032	agaaaagctctctgtctgtctcggaatcccggagctgtgtatctcgtctcaggtagtgt	1091
D	b	1104	AGAAAAGTCTCTGTCTGCTCTGGAAATCCCAGACCTTGGTATCTGTCAGGTAGTTGT	1163
Q	y	1092	aatgaatccaagaacagcacccggggtatccattccaacaatgcaaaatgtaatctgcagaacaa	1151
D	b	1164	AATGAATCCAAGAACGACCCAGGGTATACTATACAATATGCAAAATGATCTGCAGGACAA	1223
Q	y	1152	caatgtlaacagatalcggcttctcatcgtctcgtccacagcttcaacaatatcgtctaaactgt	1211
D	b	1224	CAATGTATACGATATCGGTTTTCATCGTGTTCACACAGTTCAAACATATCGGTAAACTGCT	1283
Q	y	1212	tgcctccaactgtgatacaatcgatcagaatcgtaaacgctctcctcgcacctcgtggatgctctg	1271
D	b	1284	TGCTTCCAACTGTCACATTCGTCAATATGAAACGTTCTCTCGACATCTGGTGTGCTCTTGT	1343
Q	y	1272	ggaattcatcccggttgatgacggttggtgggaacgtctccgtctgtaa	1317
D	b	1344	GGAATCATCCCGGTTGATGACGGTGGGGTAGAACGTCCGCTGTA	1389
<hr/>				
RESULT 8				
CBROTAG 4292 bp DNA linear BCF 12-SEP-1993				
LOCUS Clostridium botulinum bota gene for type A neurotoxin.				
DEFINITION X52066 X52088				
ACCESSION X52066.1 GI:40381				
VERSION bota gene; neurotoxin; secreted protein.				
KEYWORDS Clostridium botulinum.				
SOURCE Clostridium botulinum				
ORGANISM Bacillus/Clostridium group; Clostridiaceae;				
Clostridium.				
REFERENCE				
AUTHORS Minton,N.P.				
TITLE Direct Submission				
JOURNAL Submitted (08-JUN-1990) Minton N.P., PHLS Centre for Applied				
Microbiology & Research, Molecular Genetics Group, Division of				
Biotechnology, Porton Down, Salisbury SP4 0JG Wiltshire, U K				
2 (bases 1 to 4292)				
REFERENCE				
AUTHORS Thompson,D.E., Brehm,J.K., Oultrem,J.D., Swinfield,T.J.,				
Shone,C.C., Atkinson,T., Mellings,J. and Minton,N.P.				
The complete amino acid sequence of the Clostridium botulinum type				
A neurotoxin, deduced by nucleotide sequence analysis of the				
encoding gene				
Eur. J. Biochem. 189 (1), 73-81 (1990)				
FEATURES				
SOURCE				
location/Qualifiers				
1..4292				
/organism="Clostridium botulinum"				
/strain="sub sp. type A, NCTC2916"				
/db_xref="taxon:1491"				
misc_feature				
16..64				
/note="dyad symmetry"				
promoter				
19..24				

RESULT 10
 ARG00031
 LOCUS 3891 bp DNA linear PAT 04-DEC-1998
 DEFINITION Sequence 27 from patent US 5736139.
 ACCESSION AR000031
 VERSION AR000031.1 GI:3962562
 KEYWORDS
 SOURCE
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 3891)
 Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.
 Treatment of Clostridium difficile induced disease
 JOURNAL Patent: US 5736139-A 27 07-APR-1998;
 ATURES Location/Qualifiers
 source 1..3891
 BASE COUNT 1580 a 392 c 629 g 1290 t
 ORIGIN

Query Match 54.9%; Score 726.4; DB 6; Length 3891;

Best Local Similarity 72.3%; Pred. No. 1.6e-176;

Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 14 tgcctacccctcagatcacaatcaagaacatcacaatccctccatccctgaacgtgcgt 73
 DB 2588 TATCTACATTTTACTGAAATATTAAGAAATATTTATTAATCTCTATATTAAGAT 2647
 QY 74 acgaatccaatcaccgtgacgcgtctcgcgtacgcgtcccaaatccaatccgtgtc 133
 DB 2648 ATGAAAGTATCATTTAATAGACTTATAGTAGCATCAAAAATTAATATTGTAGTA 2707
 QY 134 aagtaactcgtatccgcgtacgcgaagaatccgaatccagctgttcaactcgtgaatcc 193
 DB 2708 AAGTAATTTTATCCCAATAGATATAAATCAATTCATTTTATTAAGAAATAGTA 2767
 QY 194 aaatcgaatctcctgaagaatgctatcgtatacaactcgtacgaagaactctcca 253
 DB 2768 AATTCAGGTATTTTAAAAAATGCTATGTATATATAGTATGTAGTAATTTTATTA 2827
 QY 254 cctccctcgtgacgcgtatccggaatcactccaactcctcctcgaacaatgataca 313
 DB 2828 CTAGCTTTTGTATGAAGAAATTCCTAAGTATTTTAAACAGTATTAAGTAAATGAATATA 2887
 QY 314 ccatccaactcgtgacgaagaacatctcgtgtggaagaatcctcgaactgaagtgaa 373
 DB 2888 CAAATTAATTAATGTATGAAATTAATTCAGATGGAAGATCATTAAATTAATGATGAA 2947
 QY 374 tcatctgacgtcgtcagagacactcaggaatcaaaacagcgtgtgtatccaatcactc 433
 DB 2948 TATATCGACTTTACAGAGATACAGGAATATAAACAAGAGATTTTAAATACAGTC 3007
 QY 434 agatgatacaatcctcgtacatacaatcgcgtgacgtcgttaccatccaacaatc 493
 DB 3008 AATATGTTATATATCATGATTAATAACAGATGATTTTGTAACTATCACTAAATAATA 3067
 QY 494 gctcgtataaccacaatctacatcaacgcgcgtcgtatcgacgaagaacgcatcca 553
 DB 3068 GATTAAATTAACCTTAATTTATATAATGGAAGATTAAATAGTCAAAAACCAATTTCAA 3127
 QY 554 atctgggtataccaagctctcctaatacatcgttcaaaactgacggtgtgtcgtgaca 613
 DB 3128 ATTATGAGTATATTCATCTAGTAAATATATATATGTTAAATAGATGTTGATGAGATA 3187
 QY 614 ctcaacgcgtacatcgtgatacaatcactcaatcgttgcgaagaagactgacgaagaag 673
 DB 3188 CACATAGATATTTTGGATAAATATTTTATTTTATTAAGGATTAATAAGAAAG 3247
 QY 674 aaatcaagaacgtgacgaacagtcgaatctcgtatccggaagaactctgaggtg 733

DB 3248 AATCAAAATTTATATGATTAATCAATCAAAATTCAGGTATTTTAAAAAGACTTTGGGGTG 3307
 QY 734 actaccgtcagtaacgaacaacgttactacatcgtgatactgtaacgtccgaacaatac 793
 DB 3308 ATTATTTACAAATATGATTAACCATTAATATATATGTTAAATTTATATGATCAAAATTAATAG 3367
 QY 794 ttgacgtccaacaatgtagtlatccgcgttaccatgtaacgtgaagaagtcgcgtgtctg 853
 DB 3368 TCGATGTAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3427
 QY 854 ttatgactccaacatcactcgtgaactcctccctgtaacgtgtgtaccatcactatca 913
 DB 3428 TAATGACTCAAAATTTTAAATTAATCAAGTTTGTATAGGGGACAAAATTTATATATA 3487
 QY 914 agaaatcgcgtcgtgtaacgaagaacatacgtctgcgaacaatgtaacgtgatacata 973
 DB 3488 AAAAATATGCTTCTGGAATTAAGATATATTTTGTGAAATATATGATGATGATGATGATGATG 3547
 QY 974 atgtgtagttaaagaacaagaatcaccgtctggtcgaatcgaatgctcgaatgctgtag 1033
 DB 3548 ATGTAGTAGTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 3607
 QY 1034 aaaagatctgtctcgtcgtgaatcccgagctgtgtatcgtcgtcgtgtgtgt 1093
 DB 3608 AAAAATTAATCAATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 3667
 QY 1094 tgaatccaagaacgcgcgcgtgatacactcaacaatgtaacgtcgcgaagaacaca 1153
 DB 3668 TGAAGTCAAAATATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 3727
 QY 1154 atgtgatacgtatcgttcaatcgttccacacgttccaacaatcgtcgaatcgtgtg 1213
 DB 3728 ATGGATATATATAGGCTTATAGATTTTCAATCAGTTTAAATATATATATATATATATATATAT 3787
 QY 1214 ctccaactcgtgatacgtcgtgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 1273
 DB 3788 CAAGTAATTTGTTATATAGCAATATGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 3847
 QY 1274 agtlatccgcgttgaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1317
 DB 3848 AATTTATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3891

RESULT 11
 AR169142
 LOCUS 3891 bp DNA linear PAT 17-DEC-2001
 DEFINITION Sequence 27 from patent US 6290960.
 ACCESSION AR169142
 VERSION AR169142.1 GI:17906912
 KEYWORDS
 SOURCE
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 3891)
 Kink,J.A., Thalley,B.S. and Stafford,D.C.
 Vaccine and antitoxin for the treatment of C. difficile disease
 JOURNAL Patent: US 6290960-A 27 18-SEP-2001;
 ATURES Location/Qualifiers
 source 1..3891
 BASE COUNT 1580 a 392 c 629 g 1290 t
 ORIGIN

Query Match 54.9%; Score 726.4; DB 6; Length 3891;

Best Local Similarity 72.3%; Pred. No. 1.6e-176;

Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 14 tgcctacccctcagatcacaatcaagaacatcacaatccctccatccctgaacgtgcgt 73
 DB 2588 TATCTACATTTTACTGAAATATTAAGAAATATTTATTAATCTCTATATTAAGATTAAGAT 2647
 QY 74 acgaatccaatcaccgtgacgcgtctcgcgtacgcgtcccaaatccaatccgtgtc 133

[illegible]

Accession	Sequence	Position
Dd 3788	ATGGAGTGAATTAATGGCTTTATAGCATTCATCAGTTTAATTAATAGCTAAACACTAGAG	3787
Oy 1214	cttcaactagtaaacatcgcgcagatgaacagttccctcgcagctctgggttgctcttg	1273
Dd 3788	CAATTAATTCGTATAATAGACAAATAGAAAGACTCTAGTAGCACTTTGGGTTCTCATGG	3844
Oy 1214	agttcattccgggttgatgacggttgggggtgagacgctccgctgtaa	1317
Dd 3848	AATTTATTCCTGTAGATGATGAGATGGGGAAGAAAGGCGCACTGTAA	3891

RESULT 12

LOCUS	AX036248	3891 bp	DNA	linear	PAT 16-NOV-2000
DEFINITION	Sequence 27 from Patent EP1041149.				

```

VERSION      AX036248.1  GT:11225866
KEYWORDS

```

ORGANISM synthetic construct
 artificial sequence.

AUTHORS
Kink, J. A., Firca, J. R., Padhye, N. V., Thalley, B. S., Stafford, D. C. and Williams, J. A.

disease
Patent: EP 1041149-A 27 04-OCT-2000;
JOURNAL

FEATURES	Location/Qualifiers
source	1. .3891

CDS

1. .3891

/db_xref="taxon:32630"

```

/codon_start=1
/transl_table=11

```

```

/db_xref="GI:11225867"
/translation="MQEVNKQENYKDPYGVNDIAYIKIPNVGOMQPKAFKHNKIKV

```

IGRMLTISVIRGIPFWGSTIDTELKVIDTNCINVIQPDGYSREELNVLITGPSADII
IOFECKSGHEVLNLTIRNGYGSTOYIRFSPDPTGFEESLEVDITNPLIGAGKFAATDPA

LOENEFRLYYNKEKDIASTLNKAKSVIGTPTSLQYMKMNVFEKELLSSEDTSGKFSVU
KIKEDKLYKMLTEIYLTEDNFVKEFKVLNRKITYLNEDKAVFINIIVPKYNYLYIDGENI

AND INTERFERING WITH THE CONDUCT OF THE OFFICIALS OF THE
 NDICIAVNNDLFFSPSEDNFTNDLNGEELISDTNIAEAENISLIDLQYYYLTFFNN
 DNEPENISIENTLSSDILGOLTEMPNIEFFPNKKYELDKYTMHYLBAOEFEHGKSR

STANKVLTVOITDNALSKRNEKWEDEYKY IYTNWMLAKVNT0IDLIRKMKALEN0M
DKIADITIIIPYIGPALNIGNMJYKDEFGALIFSGAIVLLEIPETAIPIVIGTFALV
RLNINYNALDEBN SNAIYIYI000I VAWY NARLLEIYD LOMVLEIPIAD I00L000000

YVONORLSTFETEXIKNIINTSLINLRSESNHLLDLSRASKINIGSKVNEFDPIDKNOY
SMIPGYKRLDEDFDASLKDALIKYIYNRGTLIGQVBRKADKQVNNNTLSTDIPEQLSK
LMIKALINTQINWILDEBRNNININIDJDSCKDNESINNEPILINAKLUNQKQSLZKIN
L

QZL NEEDS A L E Y L A N A Y L A I N C H I L E A N I S E M I N A T I F I C A S I O N N E L I I N C O M E N S
G K V S L A N G E I I W T L D P T Q E I K O R V E K S Y S M I N I S D Y I N R A F E V T I T N N R L N K S I Y
I N G R L I D O K P I S N I G N I H A S N N I M F E K D G C R Q M I N D Y I W I K Y F L F D E L N E N E I K D Y

IISALEIPDYGNLSOYVYMSKNDQGTJNKCMMNLJONNGNDJGFIHQFOFNIAKLVY
 IYLNSSLYRGTEFIILKKYASGKNDIVRNDRVYINVVYKNEYRLATNASQAVEK
 DMQSSNSGLNDEMGDITQIDAFILMLDILDFNNILVDVNNVNGIMETLNGFOSVTLI
 DMSQSSNSGLNDEMGDITQIDAFILMLDILDFNNILVDVNNVNGIMETLNGFOSVTLI

BASE COUNT	1580	a	392	c	629	g	1290	t
ORIGIN	ACGNNINQIENSRNIDGSMETPFVDDGNGEFL							

Query Match 54.9% Score 726.4 DB 6 Length 3891

best local similarity 72.5%, Freq. MO. 1.06 1/0,
Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

Db 2588 TAATCTACATTTAGTAATATATTAAAGAAATATTTAAACCTTATATTGAAATTTAAAGAT 2647

[illegible]

OY 378 ctggaactctgcaggaactcaagaaatcaaacgcgtgtgtatccaactccaat 437
|||||
DB 404 CTGGACTCTGGCAGACTGCTGTGTAACACAGAAACTGGTTTCAACTACACTCAGAT 463
|||||
OY 438 gatcaacatctcgaatcaatcaatcgcgtggaatcgttaccatcaaccaaatcgtct 497
|||||
DB 464 GATCTCTATCTCTGACTACATTAATAATGATCTTCGTTACTATCACTAACACCGTCT 523
|||||
OY 498 gaataactccaatctcaatcaacgcgcgtgtgtaacgcgaacccgatctccaatct 557
|||||
DB 524 GGGTACTCTCGTATCTCATCAACGGTAACTGATGATGAAATAATCTATCTTAACT 583
|||||
OY 558 ggttaacatccaacgtctcaatcaatcaatcgttcaaacgtgaggtgtgtgcgacatca 617
|||||
DB 584 GGGTACATCCAGCTTCTGTAACACATCTCTCAAAATCGTTGGTGGCAACGACAC--- 640
|||||
OY 618 ccgctacatctggaatcaatcaatcgttgcgacaaagaactgaacgaacgaacgaac 677
|||||
DB 641 GCGTTACGTTGGTATCCGTTACTCAAAAGTTTCGACACTGAACCTGGGTAAACTGAAT 700
|||||
OY 678 caaagactgtacgacaaacccgtlcaatctgttccgtgaagactctggtgtgacta 737
|||||
DB 701 GAAACTCTGTACTCTGAGAACCGACCGCTATCTGAAAGACTTCTGGGTAACCTA 760
|||||
OY 738 cctggaactgcagaaacccgtlcaatcgtgtgacgtgtgacgtgcgaacaaatagctga 797
|||||
DB 761 CCTGCTGTACACAAACGTTACTACTGTAACCTGCTCCGACTGACAAATCTATCAC 820
|||||
OY 798 cgtcaacaatgtagatccgcggtlcaatcgtgaacgtgaaggtccgcgtgtgtgtat 857
|||||
DB 821 TCAGAACTCT-----ACCTTCCTGAACATCAACACGACGCGGTGTTATCA 868
|||||
OY 858 gactaccaacatctcaatcgtctccgtgtacccgtgtgacaaatctcaatcaaga 917
|||||
DB 869 GAAACCTATATCTCTTACACTCGTGTGTAACCTGCTGTTGAAGTTATCATCCGTAA 928
|||||
OY 918 ataccgctct-----gtaacaagaacaatatcgttgcgaacaatgatlgtatata 968
|||||
DB 929 AAACGGTTCTACTGACATCTCATACACTGACAACTTCGTACGTAATAAACGACCTGCTTA 988
|||||
OY 969 catcaatgtttagtgaagaacaagaatccgctgtgactaccaatgcttccaggtg 1028
|||||
DB 989 CATCAACGTTGTTGACCGGTGGAATACCGTCTGTACGCTGACATCTATCCGTAA 1048
|||||
OY 1029 tctagaagaagatcgtctgcctctgaaatcccgagctgtgtaa---ctgtctcaggt 1085
|||||
DB 1049 ACCGGAATAATCATCAAACTGATCCGTACTTCTTAACACTTAACAACTCTCTGGGTAGAT 1108
|||||
OY 1086 agtltgaatgaatccaagaacgacgaaggtlactcaacaatgcaaatgaatctga 1145
|||||
DB 1109 CATCGTTATGSACTC-----GATCGGTAACTGCACTATGAACTTCCA 1153
|||||
OY 1146 ggaacaacaatggtlaacagatatcgttccatcgttccacccagttcaacaatatcgttaa 1205
|||||
DB 1154 GAACAACAACGGGTGATACATGCTGCTGGGTTTCCACTTAACAC----- 1202
|||||
OY 1206 actgtgtgtctcaactgttcaaatcgtlcaaatcgtlcaaatcgtlcaaatcgtlcaaat 1265
|||||
DB 1203 -CTGGTGTCTTCTTCTGTTACTACACAACTCCGTAAACACTTCTTCTTACGGTTG 1261
|||||
OY 1266 ctcttgggaatcaccgcgttgaatgaacgttgggtgaa 1305
|||||
DB 1262 CTTCTGTCTTCTCATCTTAAGAACACGCTGGCAGGAA 1301
|||||

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:42:55 : Search time 631.73 Seconds
(without alignments)
3595.647 Million cell updates/sec

Title: US-09-611-419A-3
Perfect score: 1323
Sequence: 1 gaattcgaaacagatgtctac.....aacgtcgctgtaagaattc 1323

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802:*

1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1323	100.0	1323	22	AA54483	Botulism toxin hea
2	1310.4	99.0	1332	22	AA54482	Botulism toxin hea
3	1310	99.0	1326	22	AA54484	Botulism toxin hea
4	1304.4	98.6	1330	17	AAT29245	Type A neurotoxin
5	1304.4	98.6	1330	19	AAV30571	Clostridium botuli
6	1304.4	98.6	1338	21	AA287212	DNA encoding synth
7	1304.4	98.6	1351	19	AAV30576	Clostridium botuli
8	1304.4	98.6	1402	17	AAT29246	Type A neurotoxin
9	1304.4	98.6	1402	19	AAV30572	Clostridium botuli

10	1250	94.5	1332	22	AA54588	Sequence encoding
11	727.6	55.0	1546	19	AAV30575	Clostridium botuli
12	727.6	55.0	4835	21	AA664582	BoNT/A neurotoxin
13	726.4	54.9	1317	21	AA287220	DNA encoding nativ
14	726.4	54.9	2532	21	AA287218	DNA encoding nativ
15	726.4	54.9	3891	17	AAT29244	C. botulinum type
16	687.4	52.0	702	21	AA287221	DNA encoding BoNTA
17	619	46.8	621	21	AA287222	DNA encoding BoNTA
18	421.6	31.9	1313	18	AAT48101	Immunogenic type F
19	421.6	31.9	1314	22	AA54499	Botulism toxin hea
20	401.2	30.3	1317	21	AA287216	DNA encoding synth
21	399	30.2	1317	22	AA54490	Botulism toxin hea
22	371.8	28.1	1917	21	AA287219	DNA encoding nativ
23	289.4	21.9	1347	21	AA287217	DNA encoding synth
24	289.4	21.9	1368	22	AA54491	Botulism toxin hea
25	264.8	20.0	1347	22	AA54589	Sequence encoding
26	263.2	19.9	1278	22	AA54489	Botulism toxin hea
27	263.2	19.9	1400	21	AA287215	DNA encoding synth
28	263.2	19.9	1400	22	AA54488	Botulism toxin hea
29	263	19.9	1341	21	AA287213	DNA encoding synth
30	263	19.9	1341	22	AA54485	Botulism toxin hea
31	227	17.2	1293	18	AAT48100	Immunogenic type F
32	203.4	15.4	1374	22	AA54487	Botulism toxin hea
33	194.2	14.7	1472	19	AAV30585	Clostridium botuli
34	194.2	14.7	4017	22	AA58862	C botulinum BoNT/E
35	191	14.4	1463	19	AAV30584	Clostridium botuli
36	180.8	13.7	1460	19	AAV30593	Clostridium botuli
37	169.4	12.8	1371	21	AA287214	DNA encoding synth
38	169.4	12.8	1371	22	AA54486	Botulism toxin hea
39	166.4	12.6	1359	12	AA012121	Synthetic tetanus
40	166.4	12.6	3712	16	AA097490	Plasmid pTRAI. N
41	166.4	12.6	3754	15	AA057880	Intermediate plasm
42	166.4	12.6	3754	16	AA085420	Plasmid pRECH1. S
43	166.4	12.6	3769	15	AA057881	Intermediate plasm
44	166.4	12.6	3769	16	AA097492	Plasmid pRECH2. N
45	166.4	12.6	4366	16	AA085424	Plasmid pRECH3-P28

ALIGNMENTS

RESULT 1	
ID	AA54483
AA54483	standard; DNA; 1323 BP.
AC	
XX	AA54483:
XX	
DT	11-APR-2001 (first entry)
XX	
DE	Botulism toxin heavy chain C-terminal coding sequence (serotype A).
XX	
KW	Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
KW	recombinant vector; antigen; Immune response; vaccine; bacterium;
KW	infection; ds.
XX	
OS	Synthetic.
OS	Clostridium botulinum.
XX	
EH	Key
FT	CDS
FT	Location/Qualifiers
FT	13..1317
FT	/tag= a
FT	/product= H_C peptide fragment
XX	
PN	MO200067700-A2.
XX	
PD	16-NOV-2000.
XX	
PF	12-MAY-2000; 2000MO-US12890.
XX	
PR	12-MAY-1999; 99US-0133865.
PR	12-MAY-1999; 99US-0133866.
PR	12-MAY-1999; 99US-0133867.
PR	12-MAY-1999; 99US-0133868.

Db 1281 ggaagtcaccccgctgcatgacgcttgaggtaacgctccgctgaagaattc 1332

RESULT 3

AAA54484
AAA54484 standard; DNA; 1326 BP.

AC AAA54484;

DT 11-APR-2001 (first entry)

DE Botulinum toxin heavy chain C-terminal coding sequence (serotype A).

KM Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
KM recombinant vector; antigen; immune response; vaccine; bacterium;
KM infection; ds.

OS Synthetic.
OS Clostridium botulinum.

Key Location/Qualifiers
FT CDS 13..1320
FT /tag= a
FT /product= H_C peptide fragment

PN WO200067700-A2.

PD 16-NOV-2000.

PF 12-MAY-2000; 2000WO-US12890.

PR 12-MAY-1999; 99US-0133865.

PR 12-MAY-1999; 99US-0133866.

PR 12-MAY-1999; 99US-0133867.

PR 12-MAY-1999; 99US-0133868.

PR 12-MAY-1999; 99US-0133869.

PR 12-MAY-1999; 99US-0133873.

PR 29-JUL-1999; 99US-0146192.

PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.

PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;

DR WPI; 2001-016048/02.

DR P-PSDB; AAB04090.

XX New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulism

PS Disclosure; Fig 3a; 73pp; English.

XX Botulinum neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.

Sequence 1326 BP; 404 A; 336 C; 243 G; 343 T; 0 other;

Query Match 99.0%; Score 1310; DB 22; Length 1326;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy	1	gaattcgaaacgagtg---ttacccttcacgtgaatacatcaagaacatcatcattctcc	57
Db	1	gaattcgaaacgagtgagcttcacgtgaatacatcaagaacatcatcattctcc	60
Qy	58	atctcgaaacctgcgtacgacatcccaatccatctgacgtctgcgtacgtctccaa	117
Db	61	atctcgaaacctgcgtacgacatcccaatccatctgacgtctgcgtacgtctccaa	120
Qy	118	atcaacatcggtcttaagaagttacattcgatccgatcgacagaatacgaatccagctgttc	177
Db	121	atcaacatcggtcttaagaagttacattcgatccgatcgacagaatacgaatccagctgttc	180
Qy	178	aatctggaatctcccaaaaacgaaagttatctctgaagaatgctatcgatatcaactatg	237
Db	181	aatctggaatctcccaaaaacgaaagttatctctgaagaatgctatcgatatcaactatg	240
Qy	238	tacgaaacctctccacactctctctgacatcgatccgatccgaataactcaactcatctc	297
Db	241	tacgaaacctctccacactctctctgacatcgatccgatccgaataactcaactcatctc	300
Qy	298	ctgaaacaatgaaatacaccaatcaatcgatcgatgaaacaatctctggttgaaagatctc	357
Db	301	ctgaaacaatgaaatacaccaatcaatcgatcgatgaaacaatctctggttgaaagatctc	360
Qy	358	ctgaactacggtgaatacatctgactctgacgagacactgagaaatcaaacagcgtgtc	417
Db	361	ctgaactacggtgaatacatctgactctgacgagacactgagaaatcaaacagcgtgtc	420
Qy	418	gtaatcaaatctctcagatgatacaaatctctgactatacaatcgctgaaatcttcgt	477
Db	421	gtaatcaaatctctcagatgatacaaatctctgactatacaatcgctgaaatcttcgt	480
Qy	478	accatcaaccaaatcgctctgaaataactccaaaatctacatcaagcgctgcgtacgcac	537
Db	481	accatcaaccaaatcgctctgaaataactccaaaatctacatcaagcgctgcgtacgcac	540
Qy	538	cagaacacgatacccaatctcgggtgaatacatccagcgtctcattacatcagttccaactg	597
Db	541	cagaacacgatacccaatctcgggtgaatacatccagcgtctcattacatcagttccaactg	600
Qy	598	gacggtgtcgtgacacatccgctacatctggaatcaaatctcaatctgttcgacaaa	657
Db	601	gacggtgtcgtgacacatccgctacatctggaatcaaatctcaatctgttcgacaaa	660
Qy	658	gaactgaaacgaaagaataccaagaacctgtagaacaacacgaatctgtgtacgtg	717
Db	661	gaactgaaacgaaagaataccaagaacctgtagaacaacacgaatctgtgtacgtg	720
Qy	718	aaagacctctggggtgactaccctgcagtagacaacacgtaactacatcgtac	777
Db	721	aaagacctctggggtgactaccctgcagtagacaacacgtaactacatcgtac	780
Qy	778	gactcgaaacaatacgtctgacgtaacaatgtagtatccggtgttaacatgtaactgaaa	837
Db	781	gactcgaaacaatacgtctgacgtaacaatgtagtatccggtgttaacatgtaactgaaa	840
Qy	838	gtctcgcggtgtctctgttatgactacaacaatctactgtgaactctctccgtacggtgt	897
Db	841	gtctcgcggtgtctctgttatgactacaacaatctactgtgaactctctccgtacggtgt	900
Qy	898	accaaatctatctcaagaatacgcgtctgtagaacaagaataatctgttgcacaat	957
Db	901	accaaatctatctcaagaatacgcgtctgtagaacaagaataatctgttgcacaat	960
Qy	958	gactcgtatatacatcaatgctgtgtagttaagaacaagaataacgctcgtgttacaatgtc	1017
Db	961	gactcgtatatacatcaatgctgtgtagttaagaacaagaataacgctcgtgttacaatgtc	1020
Qy	1018	tctcaggtcgtgtgtagaagaatctgtctgtctcgtgaatctcccgagcgttgtaactgt	1077

|||||
Db 1021 tctcagctggttagaagaatctgtctgtcgtgaatcccgagcttgtaactcg 1080
Qy 1078 tctcaggaagttagaataatccaagaacagcaggtatcactacaataatgcaaatg 1137
Db 1081 tctcaggaagttagaataatccaagaacagcaggtatcactacaataatgcaaatg 1140
Qy 1138 aatctgcgagcaacaacagtaacgatalcgtgttcacgtgttcaccacgttcaacat 1197
Db 1141 aatctgcgagcaacaacagtaacgatalcgtgttcacgtgttcaccacgttcaacat 1200
Qy 1198 atcgttaactgtgtcttccaactgtgtacaatcgttcagatcgaacgttccctcgcact 1257
Db 1201 atcgttaactgtgtcttccaactgtgtacaatcgttcagatcgaacgttccctcgcact 1260
Qy 1258 ctgggtgtctcttgaggatcattcccggtgatgacggttggtggaacgttcgctgttaa 1317
Db 1261 ctgggtgtctcttgaggatcattcccggtgatgacggttggtggaacgttcgctgttaa 1320
Qy 1318 gaattc 1323
Db 1321 gaattc 1326

RESULT 4
AAT29245
ID AAT29245 standard; DNA; 1330 BP.

AC AAT29245;

DT 07-JUL-1996 (first entry)

XX Type A neurotoxin C fragment synthetic gene.

DE Toxin: neurotoxin; fusion protein; antitoxin; vaccine; immunogen;

KW Clostridium botulinum; ds.

XX Synthetic.

XX Key Location/Qualifiers

FT CDS 1..1317

FT /*tag= a

PN WO9612802-A1.

PD 02-MAY-1996.

XX 23-OCT-1995; 95WO-US13737.

PR 07-JUN-1995; 95US-0480604.

PR 24-OCT-1994; 94US-0329154.

PR 16-MAR-1995; 95US-0405496.

PR 14-APR-1995; 95US-0422711.

XX (OPHI-) OPHIDIAN PHARM INC.

XX Firca JR, Rink JA, Padhye NV, Stafford DC, Thalley BS;

PI Williams JA;

XX WPI; 1996-230603/23.

DR P-PSDB; AAR95008.

XX Fusion proteins comprising non-toxin protein and part of toxin

PT C. difficile type toxins, and to treat C. difficile intoxication,

PT Patric. diarrhoea

XX Example 22: Page 336-38; 434p; English.

XX A synthetic gene (AAT29245) codes for the heavy chain C fragment

CC (AAR95008) of Clostridium botulinum type A neurotoxin (see also

CC AAR95010). Codon usage allowing efficient gene expression in Escherichia

CC coli was utilised. The gene in vector pAlterBot was used to make

CC expression constructs in which fragments of C. difficile toxin A
CC repeat domains were expressed as genetic fusions with the C.
CC botulin C fragment and expressed in E. coli.
XX Sequence 1330 BP; 400 A; 339 C; 246 G; 345 T; 0 other;

Query Match 98.6%; Score 1304.4; DB 17; Length 1330;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 gatgttactcactgaatacatcaagaacatcatcattcactcctgaactcgtg 71
Db 12 gctgttactcactgaatacatcaagaacatcatcattcactcctgaactcgtg 71
Qy 72 ctacgaatccaatcactcgtatcgtacctgtctcgtacgcttccaaataacatcgtgtc 131
Db 72 ctacgaatccaatcactcgtatcgtacctgtctcgtacgcttccaaataacatcgtgtc 131
Qy 132 taaagttaactcgtatcgtacgaagaatcgaatcgaatcgaatcgaatcgaatcgaatc 191
Db 132 taaagttaactcgtatcgtacgaagaatcgaatcgaatcgaatcgaatcgaatcgaatc 191
Qy 192 caaaatcgaagttaactcgtacgaagaatcgtatcgtacgaagaatcgtatcgtacgaaga 251
Db 192 caaaatcgaagttaactcgtacgaagaatcgtatcgtacgaagaatcgtatcgtacgaaga 251
Qy 252 caactcctcttgatcgcgtatccgaaataactcgaatcgaatcgaatcgaatcgaatcga 311
Db 252 caactcctcttgatcgcgtatccgaaataactcgaatcgaatcgaatcgaatcgaatcga 311
Qy 312 caactcctcttgatcgcgtatccgaaataactcgaatcgaatcgaatcgaatcgaatcga 371
Db 312 caactcctcttgatcgcgtatccgaaataactcgaatcgaatcgaatcgaatcgaatcga 371
Qy 372 aatcattcgtactcgtcgaagaactcgaagaatcgaagaactcgaagaactcgaagaact 431
Db 372 aatcattcgtactcgtcgaagaactcgaagaatcgaagaactcgaagaactcgaagaact 431
Qy 432 tgaagtatcaaatcattcgtatcgaagaactcgaagaactcgaagaactcgaagaactcga 491
Db 432 tgaagtatcaaatcattcgtatcgaagaactcgaagaactcgaagaactcgaagaactcga 491
Qy 492 tctgtgaataactcgaagaatcgaagaactcgaagaactcgaagaactcgaagaactcga 551
Db 492 tctgtgaataactcgaagaatcgaagaactcgaagaactcgaagaactcgaagaactcga 551
Qy 552 caactcgttgatcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 611
Db 552 caactcgttgatcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 611
Qy 612 caactcgttgatcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 671
Db 612 caactcgttgatcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 671
Qy 672 agaaatcaaaagcgtgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 731
Db 672 agaaatcaaaagcgtgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 731
Qy 732 tgaactcgtcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaact 791
Db 732 tgaactcgtcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaact 791
Qy 792 cgttgaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaact 851
Db 792 cgttgaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaact 851
Qy 852 tgttgaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaact 911
Db 852 tgttgaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcgaagaact 911
Qy 912 caagaatacgcgtctgtgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 971
Db 912 caagaatacgcgtctgtgaagaactcgaagaactcgaagaactcgaagaactcgaagaactcga 971

```

Db 912 caagaatacgcgtcgtgaacaagaacataatcgttcgcacaatgatcgtatatacat 971
Qy 972 caatgttgaagtaagaacaagaataacacgtctcgtctacccaatgtcttcaggctgtgt 1031
Db 972 caatgttgaagtaagaacaagaataacacgtctcgtctacccaatgtcttcaggctgtgt 1031
Qy 1032 agaaaagaatcttctcgtctcgtgaatcccgagcgttggtaactcgtctcaggtaagt 1091
Db 1032 agaaaagaatcttctcgtctcgtgaatcccgagcgttggtaactcgtctcaggtaagt 1091
Qy 1092 aatgaataccaagaacagcagggtatcactaacaatgaataatgaatgaagaaca 1151
Db 1092 aatgaataccaagaacagcagggtatcactaacaatgaataatgaatgaagaaca 1151
Qy 1152 caatggtaacgataatcgtttcattcgcacagttcaacaataatcgttaactgtt 1211
Db 1152 caatggtaacgataatcgtttcattcgcacagttcaacaataatcgttaactgtt 1211
Qy 1212 tgcctccaactgtgtacaatcgtcagatcgaaagcttcctcgcacactcgtgtgtctct 1271
Db 1212 tgcctccaactgtgtacaatcgtcagatcgaaagcttcctcgcacactcgtgtgtctct 1271
Qy 1272 ggaattcacccggttgatgaacggttgagggtgaacgttcgcgttaa 1317
Db 1272 ggaattcacccggttgatgaacggttgagggtgaacgttcgcgttaa 1317

RESULT 5
AAV30571
ID AAV30571 standard; DNA; 1330 BP.
XX
AC AAV30571;
XX
DT 07-DEC-1998 (first entry)
XX
DE Clostridium botulinum toxin A fragment C gene in palterBot.
XX
KW Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
KW botulism; ds.
XX
OS Clostridium botulinum serotype A.
XX
FH Key Location/Qualifiers
FT CDS 1..1317
FT misc_difference 1..6
FT /tag= a
FT /note= b
FT /note= "PALTER vector-derived nucleotides
FT (encode Met-Ala)"
XX
PN W09808540-A1.
XX
PD 05-MAR-1998.
XX
PF 28-AUG-1997; 97W0-US15394.
XX
PR 28-AUG-1996; 96US-0704159.
XX
PA (OPHI-) OPHIDIAN PHARM INC.
XX
PI Thalleys BS, Williams JA.
XX
DR WPI: 1998-230234/20.
DR P-PSDB; AAW68389.
XX
PT Host cell containing recombinant expression vector encoding
PT Clostridium botulinum type B or E toxin - useful to treat humans
PT and other animals at risk of intoxication with Clostridial toxin
XX
PS Example 22; Page 262-263; 428pp; English.
XX
CC This is the DNA sequence of the Clostridium botulinum serotype A
CC toxin C-fragment gene contained in plasmid palterBot. Recombinant
CC

```

```

CC C-fragment proteins have been produced in Escherichia coli as
CC fusion proteins with either maltose binding protein or
CC Clostridium difficile type A toxin (see AAW68387). The invention
CC relates to recombinant proteins derived from C. botulinum toxins.
CC Methods are provided which allow for the isolation of soluble
CC recombinant proteins free of significant endotoxin contamination.
CC Preferred hosts for production of recombinant proteins are E. coli,
CC insect cells and yeast cells. The recombinant toxin proteins are
CC used as immunogens for the production of vaccines and antitoxins
CC that are useful in the treatment of humans and animals at risk of
CC intoxication with clostridial toxin.
XX
SQ Sequence 1330 BP; 400 A; 339 C; 246 G; 345 T; 0 other;

Query Match 98.6%; Score 1304.4; DB 19; Length 1330;
Best local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 gatgtctacactgaatgaatatacaagaacatcatcaactcactcgaactgtg 71
Db 12 gatgtctacactgaatgaatatacaagaacatcatcaactcactcgaactgtg 71
Qy 72 ctacgaatccaatcactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 131
Db 72 ctacgaatccaatcactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 131
Qy 132 taaagttaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 191
Db 132 taaagttaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 191
Qy 192 caaatcgaagttaactcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 251
Db 192 caaatcgaagttaactcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 251
Qy 252 cactctcctcgtgatacgtatccgaataatcactcactcactcactcactcactcact 311
Db 252 cactctcctcgtgatacgtatccgaataatcactcactcactcactcactcactcact 311
Qy 312 caccatcatcaatcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 371
Db 312 caccatcatcaatcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 371
Qy 372 aatcatcctgactcgtcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 431
Db 372 aatcatcctgactcgtcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 431
Qy 432 tcaagatgaataacatcctgactgaatgaatgaatgaatgaatgaatgaatgaatgaat 491
Db 432 tcaagatgaataacatcctgactgaatgaatgaatgaatgaatgaatgaatgaatgaat 491
Qy 492 tgcgtgaataactccaataatcactcaatcgaatcgaatcgaatcgaatcgaatcgaat 551
Db 492 tgcgtgaataactccaataatcactcaatcgaatcgaatcgaatcgaatcgaatcgaat 551
Qy 552 caatctgggtgaatcactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 611
Db 552 caatctgggtgaatcactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 611
Qy 612 cactcaccgctacatcgtgataatgaatgaatgaatgaatgaatgaatgaatgaatgaat 671
Db 612 cactcaccgctacatcgtgataatgaatgaatgaatgaatgaatgaatgaatgaatgaat 671
Qy 672 agaatcaaaagactgtacgacaacacagttcactcgtatcctgaaagactctgggg 731
Db 672 agaatcaaaagactgtacgacaacacagttcactcgtatcctgaaagactctgggg 731
Qy 732 tgaactacgtcagtaacgacaacacagttcactcgtatcctgaaagactctgggg 791
Db 732 tgaactacgtcagtaacgacaacacagttcactcgtatcctgaaagactctgggg 791
Qy 792 cgttgacgtcaaaagttagtatacgcgggtttacatgactcctgaaggtccggtgttc 851
Db 792 cgttgacgtcaaaagttagtatacgcgggtttacatgactcctgaaggtccggtgttc 851

```

Dh 792 cgtgacgtcaacaatgtagtgcgcgtttacatgtaacgtgaaggctgcgtgttc 851
Qy 852 ttttctgctacaacaatcactgaactcttcctgtacgtgtgtaccaattcatac 911
Dh 852 ttttctgctacaacaatcactgaactcttcctgtacgtgtgtaccaattcatac 911
Qy 912 caagaatacgcgtctgtgttaacaagaacatactgttcgaacatgtatcgtatac 971
Dh 912 caagaatacgcgtctgtgttaacaagaacatactgttcgaacatgtatcgtatac 971
Qy 972 caatgtgttagttaagaacaagaatacgcgtctgttcaccaatgtcttcaggctgtgt 1031
Dh 972 caatgtgttagttaagaacaagaatacgcgtctgttcaccaatgtcttcaggctgtgt 1031
Qy 1032 agaaaagctctgtctgtctgtgaataccgcgagctgtgtatctgtctcagttgtgt 1091
Dh 1032 agaaaagctctgtctgtctgtgaataccgcgagctgtgtatctgtctcagttgtgt 1091
Qy 1092 aatgaatacgaagacgcgcgcgtgtatcactacaataatgaatgtatcgcagaca 1151
Dh 1092 aatgaatacgaagacgcgcgcgtgtatcactacaataatgaatgtatcgcagaca 1151
Qy 1152 caatgtgtatcgcgtatcgtttcattcgtttccaccagttccaacatcgttaactgtt 1211
Dh 1152 caatgtgtatcgcgtatcgtttcattcgtttccaccagttccaacatcgttaactgtt 1211
Qy 1212 tgcctcaactggtacatcgttcagatcgaacgtctcctcgtcactcgtgtgtctgt 1271
Dh 1212 tgcctcaactggtacatcgttcagatcgaacgtctcctcgtcactcgtgtgtctgt 1271
Qy 1272 ggaagtcacccgcgtgtgtatgaacgttgggtgaacgtccgcgtgtaa 1317
Dh 1272 ggaagtcacccgcgtgtgtatgaacgttgggtgaacgtccgcgtgtaa 1317

RESULT 6

AAZ87212 standard; DNA; 1338 BP.

AA87212;

08-MAY-2000 (first entry)

DNA encoding synthetic BONT serotype A (BONTA) Hc fragment.

Botulinum neurotoxin; heavy chain; BONT; serotype A;

C-terminal fragment; Venezuelan equine encephalitis virus replicon;

VEE; botulinum; vaccine; diagnosis; drug screening; ds.

Clostridium botulinum.

Synthetic.

Location/Qualifiers
9..1325
/*tag= A
/product= "Synthetic botulinum neurotoxin serotype A
(BONTA) heavy chain C-terminal fragment (Hc)"

WO200002524-A2.

20-JAN-2000.

09-JUL-1999; 99WO-US15570.

10-JUL-1998; 98US-0092416.

12-MAY-1999; 99US-0133870.

(USME-) US MEDICAL RES INST INFECTIOUS DISEASES.

Lee JS, Pushko P, Smith JF, Parker M, Dertbaugh MT, Smith L;

WPI; 2000-160827/14.

P-PSDB; AAY77134.

XX Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT toxin serotypes A-G, is used for inducing an immune response against
PT botulinum -
PS Disclosure: Page 54; 54pp; English.

The invention relates to novel vaccines that induce a protective immune response against botulinum neurotoxin (BONT) serotypes A, B, C, D, E, F and G (BONTA-BONG). The vaccine of the invention is novel recombinant CC DNA construct comprising a vector, and at least one nucleic acid fragment comprising a C-terminal heavy chain fragment (HC) from BONT serotypes A-G. In preferred embodiments of the invention, the vector is a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this vector results in the production of large amounts of a protein encoded by a sequence cloned into the replicon. The constructs are used to produce vaccines against botulism. The proteins can also be used as diagnostic tools for the diagnosis of botulism. The transformed host cells can be used to analyse the effectiveness of drugs and agents which inhibit toxin effects. The vaccine currently used against botulism is dangerous and expensive to produce, and contains formalin, which is very painful for the recipient. Also, the vaccine is incomplete, in that only 5 of the 7 serotypes are represented in the formulation. The novel CC vaccine of overcomes these problems, as it is easily purified, and available in large quantities. It is also expressed in the lymph nodes for a better immune response. Sequences AAZ87212-287217 represent synthetic DNA sequences encoding BONT Hc fragments used in the present invention. These were optimised for codon usage for expression in yeast.

Sequence 1338 BP; 401 A; 342 C; 249 G; 346 T; 0 other;

Query Match 98.6%; Score 1304.4; DB 21; Length 1338;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 gatgtctacctcaactgaatatacgaagacatcaatcactccatcgtgaactcgc 71
Dh 20 gctgtctacctcaactgaatatacgaagacatcaatcactccatcgtgaactcgc 79
Qy 72 ctaagaatcaatcacctgatcgaacctgtctgcgtacgtctccaaaatacaactcgttc 131
Dh 80 ctacgaatcaatcacctgatcgaacctgtctgcgtacgtctccaaaatacaactcgttc 139
Qy 132 taaagtaacttcgatcgcgatcgaagaatacgaatccgcgttcaactcgtgaacttc 191
Dh 140 taaagtaacttcgatcgcgatcgaagaatacgaatccgcgttcaactcgtgaacttc 199
Qy 192 caaatcgaagtatccttaagaatgctaactgaataactcgtatgaagaacttc 251
Dh 200 caaatcgaagtatccttaagaatgctaactgaataactcgtatgaagaacttc 259
Qy 252 caactccttcgtgatcgcgtatccggaataactcgaatcctcctcgaagaatgaata 311
Dh 260 caactccttcgtgatcgcgtatccggaataactcgaatcctcctcgaagaatgaata 319
Qy 312 cagcatcgaactgataatgaagaatcgtgtggaagaatcctgaagaatgaata 371
Dh 320 cagcatcgaactgataatgaagaatcgtgtggaagaatcctgaagaatgaata 379
Qy 372 aatcatctgactctcgaagacactcgaagaatacgaagcgtgtgtatctcaataactc 431
Dh 380 aatcatctgactctcgaagacactcgaagaatacgaagcgtgtgtatctcaataactc 439
Qy 432 tcaagatgatacaatcctcgaactacatcgaactcgtgatcctcgtttacaatacaaca 491
Dh 440 tcaagatgatacaatcctcgaactacatcgaactcgtgatcctcgtttacaatacaaca 499
Qy 492 tgcgtgaaataactcgaagaatacgaagaacgcgtcgtgatcgaagaagaacgcgtc 551
Dh 500 tgcgtgaaataactcgaagaatacgaagaacgcgtcgtgatcgaagaagaacgcgtc 559
Qy 552 caatcgtgtaacatccagcttctaataacatcagtccaactcgtgacgtgtgtcgtga 611

```

Db      560 caatcggggaacacccgcttctaataacatcgttaaacctgagcgttgcgtgta
Qy      612 caactacccgtatcatctgtgtaacatactcaatctgttcgacaagaactgagcaaaa
Db      620 cactcccgctaatcatctgtgtaacatactcaatctgttcgacaagaactgagcaaaa
Qy      672 agaatacaaaagaccgtgacgacaaccagttccaattctgtatctctgaagaactctggg
Db      680 agaatacaaaagaccgtgacgacaaccagttccaattctgtatctctgaagaactctggg
Qy      732 tgaactacccgtgacgacaaccagttccaattctgtatctctgaagaactctggg
Db      740 tgaactacccgtgacgacaaccagttccaattctgtatctctgaagaactctggg
Qy      792 cgttcgcttcaacaacatgtagtaaccggttcaatgtagtaaccggttcaacaacatgtagta
Db      800 cgttcgcttcaacaacatgtagtaaccggttcaatgtagtaaccggttcaacaacatgtagta
Qy      852 tgtatgactaccacaacatctactgaactcttccctgtacccgtgtgtaaccaattcatcat
Db      860 tgtatgactaccacaacatctactgaactcttccctgtacccgtgtgtaaccaattcatcat
Qy      912 caagaataacgctgtgtgtaacagaacaataatcgttgcgaacaatgtagtctgtatcat
Db      920 caagaataacgctgtgtgtaacagaacaataatcgttgcgaacaatgtagtctgtatcat
Qy      972 caatgtgtgtagtaagaacaagaataacgctgtgtgtaaccaatgtagtctgtatcat
Db      980 caatgtgtgtagtaagaacaagaataacgctgtgtgtaaccaatgtagtctgtatcat
Qy      1032 agaaaagatctgtctgtctgtgtaaacctccgagcgttgtgtaactgtctcgaagtagt
Db      1040 agaaaagatctgtctgtctgtgtaaacctccgagcgttgtgtaactgtctcgaagtagt
Qy      1092 aatgaataatccaagaagaccaggttctactaacaataatgtaattctgacgacaa
Db      1100 aatgaataatccaagaagaccaggttctactaacaataatgtaattctgacgacaa
Qy      1152 caatgttaacgatactggtttcactcgtttccacacagttcaacaataatcgttaaacgt
Db      1160 caatgttaacgatactggtttcactcgtttccacacagttcaacaataatcgttaaacgt
Qy      1212 tgcctccaactgtgtaacatgtagatcgaaagcttccctcgcacactgtggtgtctgt
Db      1220 tgcctccaactgtgtaacatgtagatcgaaagcttccctcgcacactgtggtgtctgt
Qy      1272 ggaatcattcccggtgtgtagcggttggggtgaacgtccgctgtaa
Db      1280 ggaatcattcccggtgtgtagcggttggggtgaacgtccgctgtaa

```

```

XX      05-MAR-1998.
PD      28-AUG-1997; 97WO-US15394.
PF      28-AUG-1996; 96US-0704159.
XX      (OPHI-) OPHIDIAN PHARM INC.
XX      Thalley BS, Williams JA.
XX      MPI: 1998-230234/20.
XX      P-PSDB; AAM68391.
PT      Host cell containing recombinant expression vector encoding
PT      Clostridium botulinum type B or E toxin - useful to treat humans
XX      and other animals at risk of intoxication with clostridial toxin
XX      Example 29; Page 279-281; 428pp; English.
PS      This is the DNA sequence of the Clostridium botulinum serotype A
XX      toxin C fragment gene contained in plasmid pPH18BOTA(syn). The
XX      encoded toxin A polypeptide (see AAM68391) has a histidine-tagged
XX      N-terminal extension. The vector was used to express native
XX      (i.e. non-fusion) soluble C fragment in Escherichia coli host
XX      cells. The invention relates to recombinant proteins derived from
XX      C. botulinum toxins. Methods are provided which allow for the
XX      isolation of soluble recombinant proteins free of significant
XX      endotoxin contamination. Preferred hosts for production of
XX      recombinant proteins are E. coli, insect cells and yeast cells.
XX      The recombinant toxins are used as immunogens for the production
XX      of vaccines and antitoxins that are useful in the treatment of
XX      humans and animals at risk of intoxication with clostridial toxin.
SQ      Sequence 1351 BP; 407 A; 348 C; 247 G; 349 T; 0 other:

Query Match      98.6%; Score 1304.4; DB 19; Length 1351;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      12 gatgtctactctactggaatatacacaagaacatcatcaatcaatctctctggaacgtg
Db      33 gctgtctactctactggaatatacacaagaacatcatcaatcaatctctctggaacgtg
Qy      72 ctacgaatccaatcaactgtatcgactctgtctgtctgtctgtctgtctgtctgtctgt
Db      93 ctacgaatccaatcaactgtatcgactctgtctgtctgtctgtctgtctgtctgtctgt
Qy      132 taaagttaacttcgatactcgatcgacaagaatcagatccagatcgttccaatctggaatct
Db      153 taaagttaacttcgatactcgatcgacaagaatcagatccagatcgttccaatctggaatct
Qy      192 caaatcgaagtatactctggaagaatgtatcgatatacaactcatgtagcaaaactctc
Db      213 caaatcgaagtatactctggaagaatgtatcgatatacaactcatgtagcaaaactctc
Qy      252 cactctctctgtgatacgtatccggaataactcaactcaatctctgtggaacaatgata
Db      273 cactctctctgtgatacgtatccggaataactcaactcaatctctgtggaacaatgata
Qy      312 caccatcatcaactgcatgtaaaacaacttctgtgtggaagatctctgtaactgagtgta
Db      333 caccatcatcaactgcatgtaaaacaacttctgtgtggaagatctctgtaactgagtgta
Qy      372 aatcatctgtagctctgaggaacactcaggaatacaacaacggtgtgtatcaaatctc
Db      393 aatcatctgtagctctgaggaacactcaggaatacaacaacggtgtgtatcaaatctc
Qy      432 tcaagatgatacaactctctactacatcaatcgctgtagtcttgcgttacatccacaaca
Db      453 tcaagatgatacaactctctactacatcaatcgctgtagtcttgcgttacatccacaaca

```

AAV30576
AAV30576 standard; DNA; 1351 BP.
AAV30576;
07-DEC-1998 (first entry)
Clostridium botulinum toxin A fragment C gene in pPH18BOTA(syn).
Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
botulinism; ds.
Clostridium botulinum serotype A.
Synthetic.
Key Location/Qualifiers
CDS 1..1338
FT /tag= a
FT
FT
PN WO9808540-A1.

QY 432 tcaatgatcaacaatctctgactatacaatcgctgcatctctgttaccatccaaca 491
 |||||
 DB 504 tcagatgatcaaatctctgactatacaatcgctgcatctctgttaccatccaaca 563
 QY 492 tggctggaataactccaataatctataccaacggtcgctgcatctgacccaagaactctc 551
 |||||
 DB 564 tcgctggaataactccaataatctataccaacggtcgctgcatctgacccaagaactctc 623
 QY 552 caatctggtaacatcaacacgctctataataacatcatgltcaaatctgacggtctgtctga 611
 |||||
 DB 624 caatctggtaacatcaacacgctctataataacatcatgltcaaatctgacggtctgtctga 683
 QY 612 cactcacgctatacttgatgaataactcaatctgtctcgacaaagaacttgacgaanaa 671
 |||||
 DB 684 cactcacgctatacttgatgaataactcaatctgtctcgacaaagaacttgacgaanaa 743
 QY 672 agaaatcaagaactctgtagacacacgactccaattctgtatctctgaagaactctgggg 721
 |||||
 DB 744 agaaatcaagaactctgtagacacacgactccaattctgtatctctgaagaactctgggg 803
 QY 732 tgactactctgtagacacacgactccaattctgtatctctgaagaactctgggg 791
 |||||
 DB 804 tgactactctgtagacacacgactccaattctgtatctctgaagaactctgggg 863
 QY 792 cgttgacgtcaacaatgtagtataccggttatactgtaacgtgaaaggtccgctgttc 851
 |||||
 DB 864 cgttgacgtcaacaatgtagtataccggttatactgtaacgtgaaaggtccgctgttc 923
 QY 852 tgtatgactaccacaaatctactctgactctctcctgtacgctgtagacacaaatctcat 911
 |||||
 DB 924 tgtatgactaccacaaatctactctgactctctcctgtacgctgtagacacaaatctcat 983
 QY 912 caagaataacgctctgtagacacaaatctactctgactctctcctgtacgctgtagacac 971
 |||||
 DB 984 caagaataacgctctgtagacacaaatctactctgactctctcctgtacgctgtagacac 1043
 QY 972 caatgtgtgtagttaaagaacaaagaataacgctctgactctctcctgtacgctgtagac 1031
 |||||
 DB 1044 caatgtgtgtagttaaagaacaaagaataacgctctgactctctcctgtacgctgtagac 1103
 QY 1032 agaaagaactctgctgctctgtagaatacccgagacgtctgtagaatactctcaggttagt 1091
 |||||
 DB 1104 agaaagaactctgctgctctgtagaatacccgagacgtctgtagaatactctcaggttagt 1163
 QY 1092 aatgaataccaaagaacgacaggtgatacactaaacaatgcaaatctgtagacaa 1151
 |||||
 DB 1164 aatgaataccaaagaacgacaggtgatacactaaacaatgcaaatctgtagacaa 1223
 QY 1152 caatgttaacgatactggttctgctgttccacaggttcaacaataatgctaaactggt 1211
 |||||
 DB 1224 caatgttaacgatactggttctgctgttccacaggttcaacaataatgctaaactggt 1283
 QY 1212 tgcctccaactggtacaaactgtagacggttccctcgcacactggtgtgctgtc 1271
 |||||
 DB 1284 tgcctccaactggtacaaactgtagacggttccctcgcacactggtgtgctgtc 1343
 QY 1272 gggagttacatcccggtgtagacggttgggggtgaacgtccgctgtaa 1317
 |||||
 DB 1344 ggaagttacatcccggtgtagacggttgggggtgaacgtccgctgtaa 1389
 |||||

RESULT 9

AAV30572 standard; DNA; 1402 BP.

AAV30572;

07-DEC-1998 (first entry)

Clostridium botulinum toxin A fragment C gene in phisBot.

AntiToxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;

KW botulin: ds.
 XX Clostridium botulinum serotype A.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1389
 FT /tag= a
 XX
 PN W09808540-A1.
 XX
 PD 05-MAR-1998.
 XX
 PF 28-AUG-1997; 97WO-US15394.
 XX
 PR 28-AUG-1996; 96US-0704159.
 XX
 PA (OPHI-) OPHIDIAN PHARM INC.
 XX
 PI Thalley BS, Williams JA.
 DR
 DR
 PI
 XX
 DR
 DR
 PI
 XX
 PT Host cell containing recombinant expression vector encoding
 PT Clostridium botulinum type B or E toxin - useful to treat humans
 PT and other animals at risk of intoxication with clostridial toxin
 PS Example 24; Page 265-267; 428pp; English.
 XX
 CC This is the DNA sequence of the Clostridium botulinum serotype A
 CC toxin C fragment gene contained in plasmid phisBot. The encoded
 CC toxin A polypeptide (see AAW68390) has a histidine-tagged
 CC N-terminal extension. The vector was used to express native
 CC (i.e. non-fusion) soluble C fragment in Escherichia coli host
 CC cells. The invention relates to recombinant proteins derived from
 CC Clostridium toxins. Methods are provided which allow for the
 CC isolation of soluble recombinant proteins free of significant
 CC endotoxin contamination. Preferred hosts for production of
 CC recombinant proteins are E. coli, insect cells and yeast cells.
 CC The recombinant toxins are used as immunogens for the production
 CC of vaccines and antitoxins that are useful in the treatment of
 CC humans and animals at risk of intoxication with clostridial toxin.
 XX
 SQ Sequence 1402 BP; 420 A; 360 C; 260 G; 362 T; 0 other;

Query Match 98.6%; Score 1304.4; DB 19; Length 1402;
 Best local similarity 99.9%; Pred. No. 0;
 Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 gatgtacacctcaatgaatacaatcaagaacatcatcaatccatccatgacgtg 71
 |||||
 DB 84 gctgtcactccatcagtaatacaatcaagaacatcatcaatccatccatgacgtg 143
 QY 72 ctacgaatccaaatcactgtagacgtctgctgctgctgctgctgctgctgctgctgctg 131
 |||||
 DB 144 ctacgaatccaaatcactgtagacgtctgctgctgctgctgctgctgctgctgctg 203
 QY 132 taaagttaactctgacatcgatcgacaaagaatcgatccagctgttcaatctggaactctc 191
 |||||
 DB 204 taaagttaactctgacatcgatcgacaaagaatcgatccagctgttcaatctggaactctc 263
 QY 192 caaatcgaaagttaacatcgatcgacaaagaatcgatcgatcgatcgatcgatcgatcgat 251
 |||||
 DB 264 caaatcgaaagttaacatcgatcgatcgacaaagaatcgatcgatcgatcgatcgatcgat 323
 QY 252 caactccttctgtagatcgatcgatcgacaaagaatcgatcgatcgatcgatcgatcgatcgat 311
 |||||
 DB 324 caactccttctgtagatcgatcgatcgacaaagaatcgatcgatcgatcgatcgatcgatcgat 383
 QY 312 caactccttctgtagatcgatcgatcgacaaagaatcgatcgatcgatcgatcgatcgatcgat 371
 |||||

Db 384 caccatcatcaactgcatggaanaaactctggttggaagatctctgaactacgtgta 443
QY 372 aatcatcggactctgcaggaactcaggaaatcaaacagcgtgtgtatccaatctac 431
Db 444 aatcatcggactctgcaggaactcaggaaatcaaacagcgtgtgtatccaatctac 503
QY 432 tcaagatgtaacaactctgactacatcaatcgctgtaacctgttaccatcaacaaga 491
Db 504 tcaagatgtaacaactctgactacatcaatcgctgtaacctgttaccatcaacaaga 563
QY 492 tgcgttgataatcaatccaatctataatcgaagcgcgtctgtatcgacgaagacgatctc 551
Db 564 tgcgttgataatcaatccaatctataatcgaagcgcgtctgtatcgacgaagacgatctc 623
QY 552 caacttggttaacatccacgcgtcttcaataacatggttcaaacggaagcgtgtgcgta 611
Db 624 caacttggttaacatccacgcgtcttcaataacatggttcaaacggaagcgtgtgcgta 683
QY 612 cactcacgcgtacatctgtagtcaaatctcaatctgttctgcaaaagacgtgaagaaaa 671
Db 684 cactcacgcgtacatctgtagtcaaatctcaatctgttctgcaaaagacgtgaagaaaa 743
QY 672 agaaatcaaaaagcgtgtacgacaaacagatccaatctgtgtatcctcgaaagactcttg 731
Db 744 agaaatcaaaaagcgtgtacgacaaacagatccaatctgtgtatcctcgaaagactcttg 803
QY 732 tgactactcgtgagtaacgaacacgttactacatgctgtaactgtgtaacgtccgaacata 791
Db 804 tgactactcgtgagtaacgaacacgttactacatgctgtaactgtgtaacgtccgaacata 863
QY 792 cgttgagtcgaacaaatgtagtlatcgcgcgtgtacatgtaactgtaaaaggtccgcgtgtctc 851
Db 864 cgttgagtcgaacaaatgtagtlatcgcgcgtgtacatgtaactgtaaaaggtccgcgtgtctc 923
QY 852 tgttatgactccaacatctactcgtgaactcttccctgtgtaacgtgtgtaccatctcatat 911
Db 924 tgttatgactccaacatctactcgtgaactcttccctgtgtaacgtgtgtaccatctcatat 983
QY 912 caagaataacgcgtctgtgtaacaaagacatactgtgcgaacatgtaactgtgtatcat 971
Db 984 caagaataacgcgtctgtgtaacaaagacatactgtgcgaacatgtaactgtgtatcat 1043
QY 972 caatgtgtagtgaagaaacaaagatccgtctgtgtaacaaatgcttccagcgtgtgt 1031
Db 1044 caatgtgtagtgaagaaacaaagatccgtctgtgtaacaaatgcttccagcgtgtgt 1103
QY 1032 agaaaagatctgtctgtctctggaatcccggaagctgtgtatcctgtctcaagtgtgt 1091
Db 1104 agaaaagatctgtctgtctctggaatcccggaagctgtgtatcctgtctcaagtgtgt 1163
QY 1092 aatgaataccaagaacgcgcgggtgatactatacaaatgtaaaatgtaactgtgcaagaaa 1151
Db 1164 aatgaataccaagaacgcgcgggtgatactatacaaatgtaaaatgtaactgtgcaagaaa 1223
QY 1152 caatgtagtaacgatatcggtttcatctggttccacaggttcaacaatatcgtaaacgtgt 1211
Db 1224 caatgtagtaacgatatcggtttcatctggttccacaggttcaacaatatcgtaaacgtgt 1283
QY 1212 tgccttcaactggtatcaatcgatcgaatcgatccctctgcgactctggtgtgtctgt 1271
Db 1284 tgccttcaactggtatcaatcgatcgaatcgatccctctgcgactctggtgtgtctgt 1343
QY 1272 ggaatttaccgcgtgtgtagcgcgtgtggtggaacgttcgcgcgtgtaa 1317
Db 1344 ggaatttaccgcgtgtgtagcgcgtgtggtggaacgttcgcgcgtgtaa 1389

RESULT 10
AAA54588
ID AAA54588 standard; DNA; 1332 BP.
XX
AC AAA54588;

DT 11-APR-2001 (first entry)
XX Sequence encoding botulinum toxin C fragment (serotype A).
DE Botulinum; toxin; neurotoxin; heavy chain; recombinant expression;
XX recombinant vector; antigen; immune response; vaccine; bacterium;
KM infection; ds.
XX
OS Synthetic.
OS Clostridium botulinum.
PN W020067700-A2.
PD 16-NOV-2000.
XX
PD 12-MAY-2000; 2000WO-US12890.
PE
XX 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.
PR 12-MAY-1999; 99US-0133869.
PR 12-MAY-1999; 99US-0133873.
PR 29-JUL-1999; 99US-0146192.
XX
PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX WPI; 2001-016048/02.
DR
XX New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulinism
XX
PS Example 7; Page 37; 73pp; English.
XX
CC Botulinism neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.
XX
SQ Sequence 1332 BP; 401 A; 342 C; 244 G; 344 T; 1 other:

Query Match 94.5%; Score 1250; DB 22; Length 1332;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1297; Conservative 1; Mismatches 2; Indels 6; Gaps 4;
QY 12 gatgtacactcactgatacatcaagaacatacaataactcatcactgaactgtg 71
Db 20 gctgtctacactcactgatacatcaagaacatacaataactcatcactgaactgtg 79
QY 72 ctacgaatccaatcaactgatacgtctgtctgtagcgttccaaatcaacatcggttc 131
Db 80 ctacgaatccaatcaactgatacgtctgtctgtagcgttccaaatcaacatcggttc 139
QY 132 taagtttaacttcggtcgcgtatgacaagaatgaatgaatcgaatcgttcaatctgtgaatcttc 191
Db 140 taacttaacttcggtcgcgtatgacaagaatgaatgaatcgaatcgttcaatctgtgaatcttc 199

QY 192 caaaatcgaaagtatccctgaaagaaatgctatcgatacaactctatgtaacgaaactctc 251
 |||||
 Db 200 caaaatcgaaagtatccctgaaagaaatgctatcgatacaactctatgtaacgaaactctc 259
 |||||
 QY 252 caactctctgatactcgataccgaaataactcaactcatctctctgaacataa 311
 |||||
 Db 260 caactctctgatactcgataccgaaataactcaactcatctctctgaacataa 318
 |||||
 QY 312 caaccatcatcaactgatactgaaacaaatctggtttgaaagatactctgatacagtgta 371
 |||||
 Db 319 caaccatcatcaactgatactgaaacaaatctggtttgaaagatactctgatacagtgta 378
 |||||
 QY 372 aatcatctggactctgcaaggacatcgaagaaatcaaacagcgtgtgtatcaatactc 431
 |||||
 Db 379 aatcatctggactctgcaaggacatcgaagaaatcaaacagcgtgtgtatcaatactc 438
 |||||
 QY 432 tcaagatgatacaatactctgatactcaatgctgtgatacttctgtatcaatcccaaca 491
 |||||
 Db 439 tcaagatgatacaatactctgatactcaatgctgtgatacttctgtatcaatcccaaca 498
 |||||
 QY 492 tgcctgataaactccaataatctataccaacgagcgtctgataccagaacagatctc 551
 |||||
 Db 499 tgcctgataaactccaataatctataccaacgagcgtctgataccagaacagatctc 557
 |||||
 QY 552 caatctgggtaacatccacgctcttaataacatcatggttcaaacctggacggtctgta 611
 |||||
 Db 558 caatctgggtaacatccacgctcttaataacatcatggttcaaacctggacggtctgta 616
 |||||
 QY 612 caataccgctatactgatacaatactcaatctggttcgaagaaatgaaagaa 671
 |||||
 Db 617 caataccgctatactgatacaatactcaatctggttcgaagaaatgaaagaa 676
 |||||
 QY 672 agaaatcaaaagacctgtgacacaacacagttcaaatctgtatctgaaagacctcgagg 731
 |||||
 Db 677 agaaatcaaaagacctgtgacacaacacagttcaaatctgtatctgaaagacctcgagg 736
 |||||
 QY 732 tgaactccctgacgtacgaacaaacgctatacagctgatactgttaagatccgacaata 791
 |||||
 Db 737 tgaactccctgacgtacgaacaaacgctatacagctgatactgttaagatccgacaata 796
 |||||
 QY 792 cgttcgctgaacaacatgtaagatctcgcggttatacgttacaatgaaagtcgcgtgttc 851
 |||||
 Db 797 cgttcgctgaacaacatgtaagatctcgcggttatacgttacaatgaaagtcgcgtgttc 856
 |||||
 QY 852 tgttatgataccaacatacttaactgaactcttccctgtacacgtgtgataccaatcatcat 911
 |||||
 Db 857 tgttatgataccaacatacttaactgaactcttccctgtacacgtgtgataccaatcatcat 916
 |||||
 QY 912 caagaaatacgcgtctgtaagaagaacataatcgttcgaaacaaatgatacgtgtatacat 971
 |||||
 Db 917 caagaaatacgcgtctgtaagaagaacataatcgttcgaaacaaatgatacgtgtatacat 973
 |||||
 QY 972 caatgtgtgataagaaacaaataacacgtctgctgataccaatgctctctcaagctgtgt 1031
 |||||
 Db 974 caatgtgtgataagaaacaaataacacgtctgctgataccaatgctctctcaagctgtgt 1033
 |||||
 QY 1032 agaaaagatctgtctgtctctggaatcccgagacgttgtgtaactgttctcaagtagtgt 1091
 |||||
 Db 1034 agaaaagatctgtctgtctctggaatcccgagacgttgtgtaactgttctcaagtagtgt 1093
 |||||
 QY 1092 aatgaaatccaagaacagacaggtgtatcactaaacaatgtaaatgatactgcagagcaa 1151
 |||||
 Db 1094 aatgaaatccaagaacagacaggtgtatcactaaacaatgtaaatgatactgcagagcaa 1153
 |||||
 QY 1152 caatgttaacgatacgtgttcaatcggtttccacacagttcaacaatgatactgtaactgt 1211
 |||||
 Db 1154 caatgttaacgatacgtgttcaatcggtttccacacagttcaacaatgatactgtaactgt 1213
 |||||
 QY 1212 tgcctcaactggtatacaatcgtaacatcgaaagcttccctcgacactctggtgtgtctgt 1271
 |||||
 Db 1214 tgcctcaactggtatacaatcgtaacatcgaaagcttccctcgacactctggtgtgtctgt 1273
 |||||
 QY 1272 ggaagtcatcccggtctgatagaacggttgggggtgaacgttcgcgtgttaa 1317
 |||||

Db 1274 ggaagtcatcccggtctgatacagcgttgggggtgaacgttcgcgtgttaa 1319
 |||||
 RESULT 11
 AAV30575
 ID AAV30575 standard; DNA: 1546 BP.
 AC AAV30575;
 XX
 DT 07-DEC-1998 (first entry)
 XX
 DE Clostridium botulinum toxin A fragment C gene in phisBcA.
 XX
 KW Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
 botulism; ds.
 XX
 OS Clostridium botulinum serotype A.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 108..1496
 FT /*tag= a
 XX
 PN W09808540-A1.
 XX
 PD 05-MAR-1998.
 XX
 PF 28-AUG-1997; 97WC-US15394.
 XX
 PR 28-AUG-1996; 96US-0704159.
 XX
 PA (OPH1-) OPHIDIAN PHARM INC.
 PI Thalley BS, Williams JA;
 XX
 DR WPI: 1998-230234/20.
 DR P-PSDB; AAM68390.
 XX
 PT Host cell containing recombinant expression vector encoding
 Clostridium botulinum type B or E toxin - useful to treat humans
 and other animals at risk of intoxication with clostridial toxin
 Example 28; Page 277-278; 428pp; English.
 PS
 CC This is the DNA sequence of the Clostridium botulinum serotype A
 CC toxin C fragment gene contained in plasmid phisBcA. The encoded
 CC toxin A polypeptide (see AAM68390) has a histidine-tagged
 CC N-terminal extension. The vector was used to express native
 CC (i.e. non-fusion) soluble C fragment in Escherichia coli host
 CC cells. The invention relates to recombinant proteins derived from
 CC C. botulinum toxins. Methods are provided which allow for the
 CC isolation of soluble recombinant proteins free of significant
 CC endotoxin contamination. Preferred hosts for production of
 CC recombinant proteins are E. coli, insect cells and yeast cells.
 CC The recombinant toxins are used as immunogens for the production
 CC of vaccines and antitoxins that are useful in the treatment of
 CC humans and animals at risk of intoxication with clostridial toxin.
 XX
 SO Sequence 1546 BP; 629 A; 163 C; 256 G; 498 T; 0 other:
 Query Match 55.0%; Score 727.6; DB 19; Length 1546;
 Best Local Similarity 72.2%; Pred. No. 1.2e-193;
 Matches 946; Conservative 0; Mismatches 364; Indels 0; Gaps 0;
 QY 14 tgcctcaactgataatacaagaacatataactccatccctgaactcgcgt 73
 |||||
 Db 193 tatctacattactgataatataaagaatattatatacttatactatgtaattgaat 252
 |||||
 QY 74 agaaatcaaacacccgatgacactgtctgcgaacttccaacaatcaacatcggttca 133
 |||||
 Db 253 atgaagaatcaatcaatgaacttatcagatgcatacaaaaataatatactgtagta 312
 |||||

a sequence cloned into the replicon. The constructs are used to produce vaccines against botulism. The proteins can also be used as diagnostic tools for the diagnosis of botulism. The transformed host cells can be used to analyse the effectiveness of drugs and agents which inhibit toxin effects. The vaccine currently used against botulism is dangerous and expensive to produce, and contains formalin, which is very painful for the recipient. Also, the vaccine is incomplete, in that only 5 of the 7 serotypes are represented in the formulation. The novel vaccine overcomes these problems, as it is easily purified, and available in large quantities. It is also expressed in the lymph nodes for a better immune response. The present sequence represents DNA encoding native BONTA heavy chain C-terminal fragment (hc) used in an exemplification of the present invention.

SQ Sequence 1317 BP; 560 A; 108 C; 211 G; 438 T; 0 other;

Query Match	54.9%	Score 726.4	DB 21	Length 1317
Best Local Similarity	72.3%	Pred. No. 2.3e+193		
Matches 943; Conservative	0	Mismatches 361	Indels 0	Gaps 0

[illegible]

QY	794	ttgacgtcaacaatgtatgataccgcggttaacatgtactgaaaggtccgcgttgcttg	853
Db	794	tcgaatgaataaatagtatagttataggaattatatacttaaaagggccatgaagtcag	853
QY	854	ttatgactaccacaatatctactgaaccttcctccctgtaacgctgtacccaattcaccacaa	913
Db	854	taatgaacacaacaattatcttaaatccaagtttgtatagggggaacaaattcattaaa	913
QY	914	agaaatacgcctcgtgaacaaggaacatacgtctgcgaacaatgatacgtgatatacaca	973
Db	914	aaaaatatgctctcgcgaataaagaataatatttgttgaataaataatgatacgtgatata	973
QY	974	atgttgtagtttaagaacaagaataaccgctctggtctacaaatgctctcagaagctgttag	1033
Db	974	atgtatgtagttaaaaaataaagaatatagtttagctactaatgatacacaagcgaagcgtag	1033
QY	1034	aaaagatcttgtctgcctctggaatcccggaagcttggttaattcgtctcaagtgagtgtaa	1093
Db	1034	aaaaatactcaagtcgacttgaanaatcctgtatgaagaaatccaagtcagtaagtagtaa	1093
QY	1094	tgaatcccaagaagaccaggtgatactcaacaatgcaaaatgaaatcgcgcaggaacaa	1153
Db	1094	tgaagtcacaataatgatacgaagaataacaataaagcaaatatgaaatttacaagataa	1153
QY	1154	atgtyaacgatactggtttcaatcggtttccaccagttccaacatacgcgtaaactggttg	1213
Db	1154	atgggaatgatatagcttattagattatcatcagtttaataatagataagctaaactag	1213
QY	1214	cttccaactggtataaatcgcgcagatcgaaacgttccctcctgcgaactcgggttgccttgg	1273
Db	1214	caagtaattgtgataataagacaataagaaagactagtagagcttgggttgcataagg	1273
QY	1274	agttcatccgggttgatgacggtttgggtgaacgtctgcgcgtgttaa	1317
Db	1274	aatttatccctgtagatgatagtgagatgggagaaagcccaactgttaa	1317
RESULT 14			
AAZ87218			
ID	AAZ87218	standard; DNA; 2532 BP.	
XX	AAZ87218;		
AC			
XX			
DT	08-MAY-2000	(first entry)	
XX			
DE	DNA encoding native botulinum neurotoxin serotype A (BONTA).		
XX			
XX	Botulinum neurotoxin; heavy chain; BONT; serotype A;		
KM	Venezuelan equine encephalitis virus replicon;		
KW	VEE; botulism; vaccine; diagnosis; drug screening; ds.		
XX			
OS	Clostridium botulinum.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..2532	
FT		/*tag= a	
FT		/product= "Native botulinum neurotoxin serotype A	
FT		(BONTA) heavy chain"	
FT		/transl_except= (pos:2509..2529, aa:Gly)	
XX			
PN	MOZ00002524-A2.		
XX			
PD	20-JAN-2000.		
XX			
PF	09-JUL-1999;	99WO-US15570.	
XX			
PR	10-JUL-1998;	98US-0092416.	
PR	12-MAY-1999;	99US-0133870.	
XX			
PA	(USME-) US MEDICAL RES INST INFECTIOUS DISEASES.		
XX			
PI	Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith LJ,		

PR 07-JUN-1995; 95US-0480604.
 PR 24-OCT-1994; 94US-0329154.
 PR 16-MAR-1995; 95US-0405496.
 PR 14-APR-1995; 95US-0422711.
 XX
 PA (OPHI-) OPHIDIAN PHARM INC.
 XX
 PI Firca JR, Klink JA, Padhye NV, Stafford DC, Thalley BS;
 PI Williams JA;
 DR WPI: 1996-230603/23.
 DR P-PSDB: AAR95010.
 XX
 XX Fusion proteins comprising non-toxin protein and part of toxin
 PT useful to form anti-toxins against Clostridium botulinum type A, and
 PT C. difficile type toxins, and to treat C. difficile intoxication,
 PT partic. diarrhoea
 XX
 XX Example 22; Page 344-350; 434pp; English.
 XX
 CC The Clostridium botulinum type A neurotoxin gene (AAT9244) has been
 CC cloned and sequenced. It encodes a protein (AAR95010) that is
 CC processed to form a dimer composed of a light and a heavy chain.
 CC The 50 kDa C-terminal portion of the heavy chain, or C fragment
 CC (AAR95008), was produced using a synthetic gene (AAT9245) having codon
 CC usage altered to improve expression in Escherichia coli. Fusion
 CC proteins of the type A toxin or C fragment, with e.g. maltose
 CC binding protein or polyhistidine affinity tag, are used to generate
 CC neutralising antitoxins and in vaccine compsns.
 XX
 XX Sequence 3891 BP; 1580 A; 392 C; 629 G; 1290 T; 0 other;

Query Match 54.9%; Score 726.4; DB 17; Length 3891;
 Best Local Similarity 72.3%; Pred. No. 3.9e-193;
 Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 14 tgcctacccctgaatatacatcaagaacatcatcatccctcgaacccgtcgt 73
 DB 2588 tatctacatttactgaatataatgaatattatattatctctatattgaattagat 2647
 QY 74 acgaatcaatcatcgtatgcacgtctgcgtacgtcccaaatcaatcgttcta 133
 DB 2648 atgaagaatattatgaatgaatctatcaggtatgcacaaataaatatgtagta 2707
 QY 134 aagtaactgcgtacgtacgaagaatcagatccagctgttcaatcgtgaatctcca 193
 DB 2708 aagtaaatcttgcataccaatagaataaaatcaatcaatcaatcaatcaatgaatga 2767
 QY 194 aatcggaagttaacctgaagaatgcgtatcgtatcacactatgtacgaagaatctcca 253
 DB 2768 aatcggaagttaacctgaagaatgcgtatcgtatcacactatgtacgaagaatctcca 2827
 QY 254 cctcctctgcgtatccgtatccgaatattactcaactcactctctcgtgaacaatgaata 313
 DB 2828 ctacgtcttgcgtatgaagaatctcgaatattactcaactcactctctcgtgaacaatgaata 2887
 QY 314 ccatataacacgtatgaagaatctcgtatgaagaatcgtatcgtgaactcgtgtaaa 373
 DB 2888 caataataatgtatgaagaatctcgaatattactcaactcactcactcactcactcactcact 2947
 QY 374 tcaatgcgtatcgtacgaacatcgaagaatcgaagaatcgtatgtatcgaactcgaactcga 433
 DB 2948 taatcgtatcgtacgaacatcgaagaatcgaagaatcgtatgtatcgaactcgaactcgaactcga 3007
 QY 434 agatgatacaacatcctcgtatcgaacatcgaactcgtatcgttgaactcgaactcgaactcga 493
 DB 3008 aatgataataatcgaactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcga 3067
 QY 494 gctgataataacatcgaactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcga 553
 DB 3068 gataataataacatcgaactcgaactcgaactcgaactcgaactcgaactcgaactcgaactcga 3127

QY 554 atctggtatacatccacgtctcgaatatacatatcgtatcaactcgaactcgtatcgtatga 613
 DB 3128 attagtaataatctatcgtatgaataataatgaatgaatgaatgaatgaatgaatgaatga 3187
 QY 614 ctacccgtacatctggtatcaatattcgtatcgtatcgtatcgtatcgtatcgtatcgtatcgt 673
 DB 3188 cacatagataatattgataaaatatttcaatctcttctgataaggaatgaatgaatgaatga 3247
 QY 674 aatacaagaacgtcgaacacgaatcgaatctcgtatcgtatcgtatcgtatcgtatcgtatcgt 733
 DB 3248 aatacaagaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 3307
 QY 734 actacgtcgaatcgaacacgtatcgaatcgtatcgtatcgtatcgtatcgtatcgtatcgtatcgt 793
 DB 3308 attattcaacatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 3367
 QY 794 tgaacgtcaacatcgtatgaatcgtatcgtatcgtatcgtatcgtatcgtatcgtatcgtatcgt 853
 DB 3368 tgaatgataatcgtatgaatcgtatcgtatcgtatcgtatcgtatcgtatcgtatcgtatcgt 3427
 QY 854 ttaagctacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 913
 DB 3428 taatgataacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3487
 QY 914 agaaatcgcgtcgtatgaacagaacacacacacacacacacacacacacacacacacacacacacac 973
 DB 3488 aaaaatcgtcgtcgtatgaacagaacacacacacacacacacacacacacacacacacacacacacac 3547
 QY 974 atgtgtatgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1033
 DB 3548 atgtgtatgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3607
 QY 1034 aaaaatcgtcgtcgtatgaacagaacacacacacacacacacacacacacacacacacacacacacac 1093
 DB 3608 aaaaatcgtcgtcgtatgaacagaacacacacacacacacacacacacacacacacacacacacacac 3667
 QY 1094 tgaatcgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1153
 DB 3668 tgaatcgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3727
 QY 1154 atgtgtatgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1213
 DB 3728 atgtgtatgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3787
 QY 1214 ctccacacgtatcgaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1273
 DB 3788 caagtataatcgtatgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3847
 QY 1274 agttacacacacacacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1317
 DB 3848 aatttctcgtatgaagaacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3891

Search completed: September 16, 2002, 21:43:19
 Job time: 12921 sec

• • • • •

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:32:06 ; Search time 139.08 Seconds
(without alignments)
2336.591 Million cell updates/sec

Title: US-09-611-419A-3

Perfect score: 1323
Sequence: 1 gaatcgaaacgatgtctac.....aacgtccgtgtaagaattc 1323

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Dackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
	1	1304.4	98.6	1330	1	US-08-480-604A-22
	2	1304.4	98.6	1330	2	US-08-405-496A-22
	3	1304.4	98.6	1330	4	US-08-915-136-22
	4	1304.4	98.6	1402	1	US-08-480-604A-25
	5	1304.4	98.6	1402	2	US-08-405-496A-25
	6	1304.4	98.6	1402	4	US-08-915-136-25
	7	726.4	54.9	3891	1	US-08-480-604A-27
	8	726.4	54.9	3891	2	US-08-405-496A-27
	9	726.4	54.9	3891	4	US-08-915-136-27
	10	166.4	12.6	1359	1	US-07-618-312A-3
	11	166.4	12.6	1359	1	US-08-280-228-3
	12	74.2	5.6	1359	1	US-07-618-312A-1
	13	74.2	5.6	1359	1	US-08-110-786A-7
	14	74.2	5.6	1359	1	US-08-280-228-1
	15	74.2	5.6	1858	1	US-08-668-381A-6
	16	60	4.5	5511	3	US-08-928-361B-2
	17	60	4.5	7334	3	US-08-928-361B-1
	18	57	4.3	5163	3	US-08-700-651-1
	19	57	4.3	5163	3	US-08-928-361B-4
	20	57	4.3	5318	3	US-08-700-651-2
	21	57	4.3	5318	3	US-08-928-361B-3
	22	40	3.0	1690	1	US-08-276-452A-24
	23	40	3.0	1690	2	US-08-798-744-24
	24	37.8	2.9	2277	1	US-08-676-967-2
	25	37.8	2.9	2277	1	US-08-676-974-2
	26	37.8	2.9	2277	2	US-09-098-487-2
	27	37.2	2.8	18596	4	US-09-318-448-11

28	36.8	2.8	1430	1	US-08-276-452A-25	Sequence 25, Appl
29	36.8	2.8	1430	2	US-08-798-744-25	Sequence 25, Appl
30	34.6	2.6	2094	4	US-08-714-918-87	Sequence 87, Appl
31	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
32	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
33	34.6	2.6	2094	4	US-09-266-417-87	Sequence 87, Appl
34	34.4	2.6	951	1	US-07-855-412B-2	Sequence 2, Appl1
35	34.4	2.6	951	2	US-08-308-887A-2	Sequence 2, Appl1
36	34.4	2.6	951	3	US-08-881-094-2	Sequence 2, Appl1
37	34.2	2.6	198	5	PCT-US95-10668-1	Sequence 2, Appl1
38	34.2	2.6	198	5	PCT-US95-10668-2	Sequence 2, Appl1
39	34.2	2.6	198	5	PCT-US95-10668-3	Sequence 3, Appl1
40	34.2	2.6	198	5	PCT-US95-10668-4	Sequence 3, Appl1
41	34.2	2.6	3292	1	US-07-814-964-12	Sequence 12, Appl
42	34.2	2.6	3292	1	US-08-258-442-12	Sequence 12, Appl
43	34.2	2.6	3292	1	US-08-328-809-7	Sequence 7, Appl1
44	34.2	2.6	3292	5	PCT-US92-11107-12	Sequence 12, Appl1
45	34	2.6	1666	4	US-09-134-607A-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-480-604A-22
; Sequence 22, Application US/08480604A
; Patent No. 5736139
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHYE, NISHA V.
; APPLICANT: FIRCA, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; PREVENTION OF C. DIFFICILE DISEASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,604A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422,711
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,496
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLTA, DIANE E.
; REGISTRATION NUMBER: 40,027

```

REFERENCE/DOCKET NUMBER: OPMD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1330 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1314
US-08-480-604A-22

```

```

Query Match      98.6%; Score 1304.4; DB 1; Length 1330;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 12 gatctacacctcaatgaatatacaagaacatcatcaactccatccatgaacctgcg 71
DB 12 gctgctaccttcactgaatatacaagaacatcatcaactccatccatgaacctgcg 71
QY 72 ctgaagaatccaatcaactgaatgcacgtctcgtcgaacctcccaaatcaaatcggttc 131
DB 72 ctgaagaatccaatcaactgaatgcacgtctcgtcgaacctcccaaatcaaatcggttc 131
QY 132 taagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 191
DB 132 taagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 191
QY 192 caaatacgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 251
DB 192 caaatacgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 251
QY 252 caactcccttcgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 311
DB 252 caactcccttcgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 311
QY 312 caacatcaatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 371
DB 312 caacatcaatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 371
QY 372 aatcatctgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 431
DB 372 aatcatctgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 431
QY 432 tcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 491
DB 432 tcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 491
QY 492 tgcctgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 551
DB 492 tgcctgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 551
QY 552 caatctggtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 611
DB 552 caatctggtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 611
QY 612 caatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 671
DB 612 caatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 671
QY 672 agaaatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 731
DB 672 agaaatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 731
QY 732 tgaatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 791
DB 732 tgaatcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 791

```

```

QY 792 cgttgacgtcaacaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 851
DB 792 cgttgacgtcaacaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 851
QY 852 tgtatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 911
DB 852 tgtatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 911
QY 912 caagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 971
DB 912 caagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 971
QY 972 caatgtgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1031
DB 972 caatgtgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1031
QY 1032 agaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1091
DB 1032 agaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1091
QY 1092 aatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1151
DB 1092 aatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1151
QY 1152 caatgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1211
DB 1152 caatgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1211
QY 1212 tgcctgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1271
DB 1212 tgcctgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1271
QY 1272 ggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1317
DB 1272 ggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1317

RESULT 2
US-08-405-496A-22
; Sequence 22, Application US/08405496A
; Patent No. 5919665
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, JAMES A.
; TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,496A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791

```



```

: FILING DATE: 31-OCT-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: INGOLIA, DIANE E.
: REGISTRATION NUMBER: 40, 027
: REFERENCE/DOCKET NUMBER: OPHD-01308
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1330 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1314
: US-08-405-496A-22

```

```

Query Match          98.6%; Score 1304.4; DB 2; Length 1330;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 12 gatgtcactcactgataatcaagaacatcaatcaatcctcgaactgag 71
Db 12 gctgcttacctgactgataatcaagaacatcaatcctcgaactgag 71
QY 72 ctacgaatcaatcactgactgactgctgctgctgctgctgctgctgct 131
Db 72 ctacgaatcaatcactgactgactgctgctgctgctgctgctgctgct 131
QY 132 taaagttaactcgtcgcgatcagacaagaatagatccagcgttcaactggaatcttc 191
Db 132 taaagttaactcgtcgcgatcagacaagaatagatccagcgttcaactggaatcttc 191
QY 132 taaagttaactcgtcgcgatcagacaagaatagatccagcgttcaactggaatcttc 191
Db 132 taaagttaactcgtcgcgatcagacaagaatagatccagcgttcaactggaatcttc 191
QY 192 caaaatcgaagatcctcgaagaatgctatcgatatacaactctatgtaagaactcttc 251
Db 192 caaaatcgaagatcctcgaagaatgctatcgatatacaactctatgtaagaactcttc 251
QY 252 caactcctctgagatcgatccgaaatcaactcactcactcactcactcactc 311
Db 252 caactcctctgagatcgatccgaaatcaactcactcactcactcactcactc 311
QY 312 caccatcactcactcactcactcactcactcactcactcactcactcactcactc 371
Db 312 caccatcactcactcactcactcactcactcactcactcactcactcactcactc 371
QY 372 aatcatcgtgactcgtcagagacactcagaaatcaaacagcggtgtgtatcaatcactc 431
Db 372 aatcatcgtgactcgtcagagacactcagaaatcaaacagcggtgtgtatcaatcactc 431
QY 432 tcaagatgatacaactcactcactcactcactcactcactcactcactcactcactc 491
Db 432 tcaagatgatacaactcactcactcactcactcactcactcactcactcactcactc 491
QY 492 tgcgttgatgatactcctcaaatatcaatcaatcactcactcactcactcactcactc 551
Db 492 tgcgttgatgatactcctcaaatatcaatcaatcactcactcactcactcactcactc 551
QY 552 caactcgtgatacaactcactcactcactcactcactcactcactcactcactcactc 611
Db 552 caactcgtgatacaactcactcactcactcactcactcactcactcactcactcactc 611
QY 612 cactcaccgtacatcctgatacaatcactcactcactcactcactcactcactcactcactc 671
Db 612 cactcaccgtacatcctgatacaatcactcactcactcactcactcactcactcactcactc 671
QY 672 agaatcaagaagcgtgatacaatcactcactcactcactcactcactcactcactcactc 731
Db 672 agaatcaagaagcgtgatacaatcactcactcactcactcactcactcactcactcactc 731

```

```

QY 732 tgactactcgtcagtaacaaacccgactacatgctgtaatcctgtaactcgaacaata 791
Db 732 tgactactcgtcagtaacaaacccgactacatgctgtaatcctgtaactcgaacaata 791
QY 792 cgttgacgtcaacatgtaagtaacgtgtaacgtgtaacgtgtaacgtgtaacgtgtaac 851
Db 792 cgttgacgtcaacatgtaagtaacgtgtaacgtgtaacgtgtaacgtgtaacgtgtaac 851
QY 852 tgttactactcactcactcactcactcactcactcactcactcactcactcactcactc 911
Db 852 tgttactactcactcactcactcactcactcactcactcactcactcactcactcactc 911
QY 912 caagaatcgcgtcgtgtaacaaagacataatcgttcgacaaatgatacgtgtaacat 971
Db 912 caagaatcgcgtcgtgtaacaaagacataatcgttcgacaaatgatacgtgtaacat 971
QY 972 caatgltgagttaagaacaaagataacccgtcgtgtaacaaatgatacgtgtaacat 1031
Db 972 caatgltgagttaagaacaaagataacccgtcgtgtaacaaatgatacgtgtaacat 1031
QY 1032 agaaaagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1091
Db 1032 agaaaagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1091
QY 1092 aatgaatccaaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1151
Db 1092 aatgaatccaaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1151
QY 1152 caatgtaacagatcgtgtaacgtgtaacgtgtaacgtgtaacgtgtaacgtgtaacgtgta 1211
Db 1152 caatgtaacagatcgtgtaacgtgtaacgtgtaacgtgtaacgtgtaacgtgtaacgtgta 1211
QY 1212 tgcctcactcactcactcactcactcactcactcactcactcactcactcactcactcactc 1271
Db 1212 tgcctcactcactcactcactcactcactcactcactcactcactcactcactcactcactc 1271
QY 1272 ggaagttcattccggttgatgacagcgttggtggtggtggtggtggtggtggtggtggt 1317
Db 1272 ggaagttcattccggttgatgacagcgttggtggtggtggtggtggtggtggtggtggt 1317

```

```

RESULT 3
US-08-915-136-22
: Sequence 22, Application US/08915136
: Patent No. 6290960
: GENERAL INFORMATION:
: APPLICANT: KINK, JOHN A.
: APPLICANT: THALLEY, BRUCE S.
: APPLICANT: PADHYE, NISHA V.
: APPLICANT: FIRCA, JOSEPH R.
: APPLICANT: STAFFORD, DOUGLAS C.
: TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
: PREVENTION OF C. DIFFICILE DISEASE
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MEDLEN & CARROLL, LLP
: STREET: 220 MONTGOMERY STREET, SUITE 2200
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/915,136
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/480,604
: FILING DATE:

```

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER: US 08/405,496
3      FILING DATE: 16-MAR-1995
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER: US 08/329,154
6      FILING DATE: 25-OCT-1994
7      PRIOR APPLICATION DATA:
8      APPLICATION NUMBER: US 08/161,907
9      FILING DATE: 02-DEC-1993
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER: US 07/985,321
12     FILING DATE: 04-DEC-1992
13     PRIOR APPLICATION DATA:
14     APPLICATION NUMBER: US 07/429,791
15     FILING DATE: 31-OCT-1989
16     ATTORNEY/AGENT INFORMATION:
17     NAME: INGOLIA, DIANE E.
18     REGISTRATION NUMBER: 40,027
19     REFERENCE/DOCKET NUMBER: OPHD-01763
20     TELECOMMUNICATION INFORMATION:
21     TELEPHONE: (415) 705-8410
22     TELEFAX: (415) 397-8338
23     INFORMATION FOR SEQ ID NO: 22:
24     SEQUENCE CHARACTERISTICS:
25     LENGTH: 1330 base pairs
26     TYPE: nucleic acid
27     STRANDEDNESS: double
28     TOPOLOGY: linear
29     MOLECULE TYPE: DNA (genomic)
30     FEATURE:
31     NAME/KEY: CDS
32     LOCATION: 1..1314
33     US-08-915-136-22

```

Query Match	Best Local Similarity	98.6%	Score 1304.4	DB 4	Length 1330
Matches 1305	Conservative	0	Mismatches	1	Indels
					Gaps
QY	12	gatgtctactccttcacatgatacatcaagaagacatcaataactccatctcctgaactgcyg	71		
Db	12	GCTGCTACCTTCTCAGTGAATATACATCAAGAACATCATCAATACCTCATCTGACCTGG	71		
QY	72	ctacgaatcccaatcaacctgattccgacctgtctcgtctacgctttccaaaatcaaatcgtgttc	131		
Db	72	CTACGAATCCCAATCAACCTGATCGACCTGCTCTCGTACCGCTTCCAAATCAATCATCGGTTTC	131		
QY	132	taaaagttaacttcgataccgatcgacaagaatcagatccacgctgttcaatctgaaatcttc	191		
Db	132	TAAAGTTAACTTCGATCCGATGACAGAGATGAGATCCAGCTGTTCAATCTGGAAATCTTC	191		
QY	192	caaaaatcgaaagttatccctgaaagaatgatactglatacaactctatgttaccgaaactcttc	251		
Db	192	CAAAAATCGAAGTTATCCCTGAAAGAAATGCTATCTGTATACAACTTATGTAGAGAAACTTCTTC	251		
QY	252	caaccctcttcggaatccgataccggaataacttcaactcaatccatctctcgaacaatgaaata	311		
Db	252	CACCTCTCTTGGATTCGGATATCCCAAAATACTTCAACTCCANTCTCTGTAACAATGAAATA	311		
QY	312	caaccatcatcaactgcatgynaanaaactctgtgttgynaagratctctgaactacgylga	371		
Db	312	CACCATCATCAACTGCATGGAAGAAACAATCTCTGTTGGAAGATATCTCTGAACTACGGTGA	371		
QY	372	aatacttcgagctctcgagggagacccaagaaatcaaaacagcgtgtgttataatcaatctc	431		
Db	372	AATCTTCGAGCTCTCGAGGACACTGAGAAATCAAAACGCTGTGTATTTCAAAATATCTC	431		
QY	432	tcaagatgatacaaatctctgatactcaatcaatcgcctggaatctcgttatcatcacaaca	491		
Db	432	TCAGATGATCAACATCTCTGACTACATCAATCGTGAGATCTCTGTATTCATCACCACAA	491		
QY	492	tcgcttgataaactccaatactatacaagcgccgcttgatcgacaggaacgpatctc	551		

Dd	492	TGCGTGAATTAACGCCAAATVCTACATCAACAGGCSCGTGTGATCGACACGAAACCATCTC	551
Oy	552	caatctggttaaacatccacgcttctaatacatcagttccaacgcggagtgcgga	611
Dd	552	CAATCTGGGTAAACAATCCACCCTTTCTAATAACATCATATGTTCAAACGTGGACGGTGTCGGA	611
Oy	612	cactcacgcgtacaactcgtgaatcaaatcactcaactcgttcgcgaagaagacttgaaagaaa	671
Dd	612	CACGTACCGGCTACGATCTGTGATFCAAAATACCTCAATCTGTTGCAACAAAAGAATCAACGAAAA	671
Oy	672	agaatacaaagacctgttagacaacaaccagtlccaattcttgttatccctggaaggacttctggg	731
Dd	672	AGAAATCAAAACACTGTACGACAAACACAGTCCAATTCTGGTATCTGTAAAGACTTTGCGGG	731
Oy	732	tgaactacctgtagcagcaacaacogtaactacatgctgaactctgtatcgatcgcgaacanaa	791
Dd	732	TGACTACTCTGCAGTAGACCAACAACCGTACTACTCATGCTGATCTGTAGATCCGACAAATA	791
Oy	792	cgttgacgticaaacaatglaagtatccgcggtttaacatgatccctgaaagtgccgcgtgttc	851
Dd	792	CGTTGACGCTCAACCAATTAGGTATCCGCGGTTACATGTACCTGAAAGGATCCCGCGTTC	851
Oy	852	tggtatgactaccaactctactctggaacttcccctgtaactctgtgtagccaattcaat	911
Dd	852	TGTTATGTACTACCAACATCTACCTGGAACCTTCCCTGTACCGGGATCCAAATFTACAT	911
Oy	912	caagaataacgcgcgtctgttaaacaagaacaatalcgtctgcaacaatgualcgtgtatatcat	971
Dd	912	CAAGAATAACGCGCTGTGTAACAAGCACATATCGTTCCAGCAACAAATGATCGTATPACAT	971
Oy	972	caatgttgttagtaagaacaagaataacgctctggtaccatagcttctcaagccgtggt	1031
Dd	972	CAATGTTGTACTTAAAGAACAAAGAAATACCTCTGGGTACCATGCTCTCAGCGTGGGT	1031
Oy	1032	agaaaagatctctgctcgtctctggaatcccgagcgttgttaatcgtctcagaatgtgt	1091
Dd	1032	AGAAAAGATCTTGTCTGTCTCTGTGAATCCCGAGCGTTGGTANTGTGTCTCAGAGTGTGT	1091
Oy	1092	aatgaatatcaaagaacgacagggatcacataaiaaycaaaatgaatctgcagagaca	1151
Dd	1092	AATGAATCCAAAGAACACACACAGGGTATCACTAACAAATGCAAAATGTCAATCTCCAGACAA	1151
Oy	1152	caatgtaacgatatcggtttcatcggttccacacagttcaacaatatcgtcaaatgt	1211
Dd	1152	CAATGTATACGATATTCGGTTTTCATCGTTTTCCACGATTCACAAATATGTGCTAAATCGT	1211
Oy	1212	tgcttccaactgtagacaatcgttcagalcgaacgltccctctgcacactcgtgtgtctctg	1271
Dd	1212	TGCTTCCAATGTTACATGATGTGTCAGATCGAAGCTTCTCTCGCAGCTCTGGTGTGCTTGG	1271
Oy	1272	ggaagtcatccccgttgtatgaaggtttgggttgtaagtlccgcgtgtaa	1317
Dd	1272	GGAGTTCATCCCGGTTGATGACGGTGGGTGGAACGTCCGCTGTAA	1317
RESULT 4			
US-08-480-604A-25			
Sequence 25, Application US/08480604A			
Patent No. 5736139			
GENERAL INFORMATION:			
APPLICANT: KINK, JOHN A.			
APPLICANT: THALLEY, BRUCE S.			
APPLICANT: PADHAY, NISHA V.			
APPLICANT: FIRICA, JOSEPH R.			
APPLICANT: STAFFORD, DOUGLAS C.			
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND			
NUMBER OF SEQUENCES: 32			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: MEDLEN & CARROLL, LLP			
STREET: 220 MONTGOMERY STREET, SUITE 2200			
CITY: SAN FRANCISCO			
STATE: CALIFORNIA			

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1386
US-08-480-604A-25

```

```

Query Match          98.6%; Score 1304.4; DB 1; Length 1402;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 312 caacatcatcaactgcatgaaacaattctggttgaaagtatctcgaactgagtg 371
DB 384 CACCATCATCAACTGCGATGGAAAAACAAATTCGTGGTGGAAAGATCTCTCACTACGGTGA 443
QY 372 aatcatctgactctgcaagacactcgaagaaatcaaacggtgtgtatccaatctc 431
DB 444 AATCATCTGACTCTGCAAGACACTCAGGAATCAAAACGCGTGTATTCGAAATATCTC 503
QY 432 taagatgatacaatctctgactataatcgaatgctggtgactctcggttaccataccaacaa 491
DB 504 TCAGATGATCAACATCTCTGACTACATCAATGCTGGATCTTCTGTTACCATATCAACAA 563
QY 492 tcgctcgaataactccaatatctacatacaagcgctcgcgactcgaccagaacacgctc 551
DB 564 TCGTCTGAATTAATCTCAAAATCTACTACATCAACGCGCTCTGATCGACCAAAACGATCTC 623
QY 552 caatctggglaacatccacgcttctaataacalcatgltcaaacctgagcgtgtgtcgtga 611
DB 624 CAATCGGGTAAACATCCACGCTTCTAATAACATCAATGTTCAAACTGACGCGTGTCTGTA 683
QY 612 caatcaccgctacatctgataccaatctcaatctgttcggaacaaagaaactgaacgaaaa 671
DB 684 CACTCACCGCTACATCTGATGATCAAAATATTCATATCTGTTGACAAAGAACTGAACGAAAA 743
QY 672 agaaatcaaaagcctgtaagacaaccagtcacaattctgtatccctgaagactctgag 731
DB 744 AGAAATCAAGACCTGTACAGACACACCAATTCATTTGGTATTCCTGAAGACTCTCGGG 803
QY 732 tgactacctcgagtaagacaacacgtaactacatgctgaaatctgtacgactcgaaacaata 791
DB 804 TGACTACCTGCGATAGCAAAACCGTATACATGCTGATCTGTACGATCCGAACAAATA 863
QY 792 cgttgaagcacaagatgtagtataccggttgaatgataatgtaacccgaaggtccggtgttc 851
DB 864 CGTTGACGTAAACATGTAGTATCCGCGTTACATGTACCTGAAAGGTCCGCGTGTGT 923
QY 852 tgtatgactacaacatctacatctcctcctcgtacacgtgtgtacccaatitacat 911
DB 924 TGTATGACTACCAACATCTACTACTGAACTTCCCTGTACCGGTGATCCAAATTCATCAT 983
QY 912 caagaatacgcgctgtgttaacaaggaacatatcgltcgacaacatgatactggtatcat 971
DB 984 CAGAATATACGCGTCTGTGTACAGACAAATATCTGTCACAAATGATCGGTATATACAT 1043
QY 972 caatgttgaagttaagaacaaagaaataacgcttgcgtacaaatgcttcgaagctgtgt 1031
DB 1044 CAAATGTGTAGTTAAGAACAAAGAAATACCGTCTGGCTACCAATGCTTCTAGGCTGTGT 1103
QY 1032 agaaagatctgtctgctctggaatcccggaacgttgttaatctgtctcaagtgatgt 1091
DB 1104 AGAAAGATCTGTGTCTGTCTGGAATTCGCCGACGTTGGTAATCTGTCTCAGGTACTTGT 1163
QY 1092 aatgaatccaaagaacgaccaggtgtatcaactaaatgcaaaatgaatctcgaagaa 1151
DB 1164 AATGAATCCAAAGAACGACGAGGTATCACTAACAATCAAAATGAATCTGAGGACAA 1223
QY 1152 caatgttaacgatatcggttcaatcggttccaacaggttcaacaatatcgcttaaacgtgt 1211
DB 1224 CAATGTATACGATATGCGTTTCAATCGGTTTCCACCGCTTCAACAAATATGCTTAATCGGT 1283
QY 1212 tgcctcacaactgtataacatcgctgaatcgaacgttccctcgcactctggtgtgtctgt 1271
DB 1284 TGCTTCAACGTGTATCAATTCGATCAGATTCGACAGTTCCTCTCGGACCTCGGGTGTCTTG 1343
QY 1272 ggaagtcatcccggtgtgatagaaggttgggtggaagtcgcgtgttaa 1317
DB 1344 GGAGTTCATCCGCGTGTGATGACGCTTGGGGTGAAGCTCCGCTGTA 1389

```

```

RESUME 5
US-08-405-496A-25
; Sequence 25, Application US/08405496A
; Patent No. 5919665

```

```

GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1386
US-08-405-496A-25

Query Match          98.6%; Score 1304.4; DB 2; Length 1402;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      12 gatgtctacattcatgatatacatcaaaaacaatcctaacttcattgaactcg 71
        | |||||
DB       84 gctgttaccttccatgtaatatcatgaagaaactatcatattaccctcatcgaacctggc 143

OY      72 ctacgaatccaatacacccgatgcagctgtctcgtcagcgtctccaatatcaaatcggttc 131
        |||||||
DB     144 CTAGGAATCCAAATCAATCACCGATGCAGACTGTCTCTGCTACGGTTCCAAAATCAACATCGGTTTC 203

OY     132 taaagttaacttgatcgatcgatcgacagaagaatagatccaagctgtttcaatctgyaatcttc 191
        |||||||
DB     204 TAAAGTTAACCTTCGATCCGATCGACACAAGAATCGATCCAGCTGTTCATCTGGAATCTTTC 263

OY     192 caaaatcgaagttatcccgaagaatgctatcgatatcaactctagttagaanaactcttc 251
        |||||||
DB     264 CAAAATCGAAGTTATCCCgaagaatgctatcgatatcaactctagttagaanaactcttc 323

```

QY	252	caaccccttcgagtcgcgtatcccggaataacttcaatccatccatctctctgaaacaagaa	311
Db	324	CACCTCTTTCGATCCGTAATCCCGAAATATCTTCAACTCCATCTCTGTGAAACAATGA	383
QY	312	caccatcatcaactgtgataagaaacaattctggttggaaagatattctctgaaactggtga	371
Db	384	CACCATCATCAACGTCATGTGAAAAACAATTCCTGGTTTGAAAGATATCTCTGMACTACG	443
QY	372	aatacttggaccttggagaacactgaagaaatcaaaacggtgtgtgtgtatcaaatctac	431
Db	444	AATATCTGGACCTCTGCAGACACTGAAGNAATCAACACGCTGTGTGTATTAATTA	503
QY	432	tcgaatgatcaaatctctctgtacataccaactcgaactgtgaaacttcgtttacatccaaca	491
Db	504	TCAATGATCAACATCTCTGATCAATCAATGCTGATCTTCTGTACCATCAACACA	563
QY	492	tcgtctgataaatactcaaaaatactacatacaacggtcgtctgtatcgaccagaacgta	551
Db	564	TCTGCTCAAAATACCTCCAAATTCATCAACGGCCCTGTGATGACCAAAACCCATCTC	623
QY	552	caatctgggtgaacataccaacgtctctaatcaatcatggttccaacggagcgtgtctgtga	611
Db	624	CAATCTGGGTACATCCACGCTTCTTAATACATCATAGTTCAAACTGGACGTTGTCTGTA	683
QY	612	cactcaacgcgttacaatctctgtataactccaactcgtctgttcgaacaaagaaactgaa	671
Db	684	CACTCACCGCTACATCTGATCAAAATACATTCATCTGTCGACAAAGAACTCAACGA	743
QY	672	agaaataaagaccttgcagacaacaacgaagccaattctgtgtatctcgtgaagaaactt	731
Db	744	AGAAATCAAAAGCCTGTGATGACACAACCAATTCATTCCTGTATCTGTGAAGACTT	803
QY	732	tgaataccttgcagtaacgaacaacccgttacaatctgtacgttcgaatctgtatcgata	791
Db	804	TGACTACTCTGAGATGACGAACAACCGTATCATAGTGTGAATGTGTAGATCCGAACA	863
QY	792	cgttgcagctacaagaatgataagtatccgcggtttacatgtatccctggaagatccgcgt	851
Db	864	CGTTGACGTCAACAAATGTAGGTATCCGCGGTATCAATGTAAGTGAAGGTCCGCTG	923
QY	852	tgattatgaataccaacaatactgaactttccctgtacacgttgcgtgtgataccaatcat	911
Db	924	TGTTATGATACCAACATCTACCTGAACCTTCCGTACCGGTGACCAATATCATAT	983
QY	912	caagaaatacgcgtctctgtatacaagaacaaatctgtctgcgaacaatgatctgtatacat	971
Db	984	CAAAATAATACCGCTCTGTGTAAAGGACAATATCTGTCGAACAAGATCCGTGTATAT	1043
QY	972	caatgtgtgattgaagaacaagaataacgctctgcgtataccaatgtctctcgaagctgt	1031
Db	1044	CAATGTGTGATTAGAACAAAGAATATCCGCTGCGCTACCAATGCTTCTCAGCTGGTGT	1103
QY	1032	agaaaagatcttgtctgtctctgtgaaatcccggaacglttggtaatctgtctcagtagt	1091
Db	1104	AGAAAATCATCTTGTCTCTCTGTGAAATCCCGACGTTGTGTAATCTGTCTCAAGT	1163
QY	1092	aatgaataccaagaacgacaggttatctactaaacaatgcaaaatgaaatctgtcagaaca	1151
Db	1164	AATGAATAATCAAGAACACCCAGGGTATCATCAATAATCAAAATGAATCTTCCAGGACA	1223
QY	1152	caatcgttaacgatactcgtttcaatctgtttccaacacggttccaacaatatcgtcacaact	1211
Db	1224	CAATCTGATGAGATATCGGTTTCAATTCGTTTCCACAGTTCAACAAATATATGCTAACT	1283
QY	1212	tgcttccaactgtgataaactcgttcagatcgaaacglttcctctcgcactcgtgtgtctgt	1271
Db	1284	TGCTTCCAACTGTGATCAATGTCAGATCGAACCTTCTCTCCACACTCTGGGTTGTCTT	1343
QY	1272	ggaattatcccggttgaataaggttgggggtgaagcttcgcgtgtaa	1317
Db	1344	GGAATTCATCCCGGTTGATATACGGTTGGGGTGAACTCCGCTGTAA	1389

```

RESULT 6
US-08-915-136-25
: Sequence 25, Application US/08915136
: Patent No. 6290960
: GENERAL INFORMATION:
: APPLICANT: KINK, JOHN A.
: APPLICANT: THALLEY, BRUCE S.
: APPLICANT: PADHYE, NISHA V.
: APPLICANT: FIRCH, JOSEPH R.
: APPLICANT: STAEFORD, DOUGLAS C.
: TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
: NUMBER OF SEQUENCES: 32
: PREVENTION OF C. DIFFICILE DISEASE
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MEDLEN & CARROLL, LLP
: STREET: 220 MONTGOMERY STREET, SUITE 2200
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/915,136
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/480,604
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/405,496
: FILING DATE: 16-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/329,154
: FILING DATE: 25-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/161,907
: FILING DATE: 02-DEC-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/985,321
: FILING DATE: 04-DEC-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/429,791
: FILING DATE: 31-OCT-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: INGOLIA, DIANE E.
: REGISTRATION NUMBER: 40,027
: REFERENCE/DOCKET NUMBER: OPHD-01763
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 25:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1402 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1386
: US-08-915-136-25
:
: Query Match 98.6%; Score 1304.4; DB 4; Length 1402;
: Best Local Similarity 99.9%; Pred. No. 0;
: Matches 1305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

D	84	GCTGTCTACCTTCAGTAATACATCAAGACATCATCAATACCTTCATCTGAAACCTGCG	143
Q	72	ctacgataccaatcacctgatcgaacctgtcgtctacgctcttccaaataccaatcagcttgc	131
D	144	CTAGGATATCAATCACCTGATGAGACCTGCTGCTACGCTTCCAAATATCAATGATGGTTC	203
Q	132	taaatgtaacttcgatccgatccgacaagaatcaagatccagctcgttcaatctcgtgaatcttc	191
D	204	TAAAGTTAACTTCGATCCGATCGACACAGATATGATCCAGCTGTTCAATCTGGAAATCTTC	263
Q	192	caaaatcgaagttatccgttaaagaatgctctgttatcaaatctatgtaacgaanaattcttc	251
D	264	CAAAATCGAAGTTATCTCTGTAAGATGCTATGATTAACAATTATATGACGAAACTTCTC	323
Q	252	caactcccttcgtgatccgtatccccgaataacttcaactccatctctctcgtgaacaatgaata	311
D	324	CACCTCCTTGTGATCCGATATCCCGAAATATCTCAACTCCATCTCTGAAACAATGATTA	383
Q	312	caaccatcaactcgtcgtatggaaaaaattctcgtgttggaagtatctctgaactcagctga	371
D	384	CACCATCTATCACTGCATGGAAACAAATCTGTTGGAAAGATCTCTGCAACTACGAGTGA	443
Q	372	aatcattcgtgacctcgtcaggaacactgaagaaatcaaacagcgctgtctatccaatctactc	431
D	444	AATCATCTGGACCTCTGAGGACACTAGGAATATCAAAACGCGTGTGTTATCAAAATCTC	503
Q	432	tcagatgatcaaatctctgtaactataccaatcgtcgtgatcttcgttatccatccaacaata	491
D	504	TCAGATGATCAACATCTCTGACTACATCAATCGCTGGATCTTCTGTACATCACCACAA	563
Q	492	tcgtctgataactccaataatctataccaacagcgctcgtatccgacccaagaacagcatctc	551
D	564	TCTGCTGAAATAACTCCAAATCTACATCTCAACGCGCTGATCGACACGAAACCATCTC	623
Q	552	caacttcgggtgaatatccacgctctctaaagcaaatggttcaaacctggcggtgtgtcgtga	611
D	624	CAATCTGGGTAAATCACGCTTCTATTAACATCATCATGTTCAAACTGAGAGGATGTGCTGA	683
Q	612	cactcaacggtataatcttgatccaataacttcaactcgtctcgaacaagaacttgaacgaata	671
D	684	CACCTACCGCTACATCTGGATCAAAATCTTCATCTGTTGCAACAAAGAACTGAACGAA	743
Q	672	agaaatcaaaagaccgttaegacaacacgattccaattctgttatcctcgtgaagactctcgtgg	731
D	744	AGAAATCAAAAGACCTGTACGACACAAACGATCCAAATCTGCTATCTCGAAAGAGACTTCTGGGG	803
Q	732	tgactctcgtcagtaacgaacaaacgactaagctgtcgaatctgtatcgatccgaacaata	791
D	804	TGACTACTCTGCAATAGCACAAACCGTATACATAGCTGAATCTGTACGATCCCAACAAATA	863
Q	792	cgttcgactcaacaatctgaagtaatcccggtttacatgtaacctggaagaagctccgctggttc	851
D	864	CGTTGACGTAACAATGTAGTATCCGCGGTTACATGTACTCGAANAAGTCCCGCTGGTTC	923
Q	852	tgatagactaccaaatcttaccctgaactcttcctctgtaacgcttggtaacaaatctaatcat	911
D	924	TGTTATGACTACCAACATGTACTGCTGAACGCTTCCTCCGTGACCGTGAATTCATCAT	983
Q	912	caagaataacacgctctgtgtaacaaggaacaatacgttcgcaacaatgactcgttatcat	971
D	984	CAAGAATATACGCTCTGTGTACACAGCAATATTCGTTCCCAACATGATCGTGTATCAT	1043
Q	972	caatgtttgtaagtaaaaaaataacgcttgcctcacaatgcttccagagcgtggt	1031
D	1044	CAATGTGTATGTTAAGAACAAAGAAATACCGCTCTGGCTACCAATGCTTCTCCAGCTGGTGT	1103
Q	1032	agaaagaatctctcgtcctcgtgnaatcccgagacgtcttgtaactcgtctcaggtatggtt	1091
D	1104	AGAAAGATCTTGTCTGCTCTGGAATTCGCCGAGCTGTGTAATCTGCTCAAGTATGTTGT	1163
Q	1092	aatgnaatccaagaacgaccaggtgtaactaatacaatgnaaatgnaatctcgtcaggaaca	1151
D	1164	AATGAAATCCAAAGACGACGAGGTATCTACTAACAATGCCAAATGCCAAATCTTCACAGACAA	1223

QY 1152 caatgtaacgatcgttccatccagttccacacatcgaactggt 1211
DB 1224 CATGTAAGATGATGTTTCCATGCTTCCAGATCAATATCGTAACTGGT 1283
QY 1212 tgcctcaactggtacacatcgtcgaacggtcctcgcactcgtggtctctg 1271
DB 1284 TGGTTCACACTGGTACATGTCAGATCGACGTTCTCTCGCAGCTGGGTCTCTTG 1343
QY 1272 ggaatcaccggttgatgacggttgaggagacgtccgctgtaa 1317
DB 1344 GGAGTTATCCCGTTGATGACGTTGGGGTGAACGTCCGCTGTAA 1389

RESULT 7
US-08-480-604A-27
Sequence 27, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3891 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3888
US-08-480-604A-27

Query Match 54.9%; Score 726.4; DB 1; Length 3891;
Best Local Similarity 72.3%; Pred. No. 1.3e-203;
Matches 943; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 14 tgcctcactcactgatacatcaagaacatcatalcactccactcgaactcgcgt 73
DB 2588 TATCTACATTTACTGATATATATTAAGAAATTTATTAATCTCTATATGAATTAAGAT 2647
QY 74 agcaatccaactcaactcgtacgactcgtcgtcgaactcccaaatccaactcgtctca 133
DB 2648 ATGAAAGTAAATCAATTAATGACTTATCTAGTATCATCAAAAATAATATTTGGTAGTA 2707
QY 134 aagtaactcgaatccgacatcgaagaatcgaatccgactcgtcgaactcgtcga 193
DB 2708 AAGTAAATTTTGCATCCAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2767
QY 194 aaatcgaagttacatcctgaagaatgctatcgtatcacaactatgtaagaactccca 253
DB 2768 AAATGAGGTAATTTTAAAAAATGCTATGTTATATATGATGATGATGATGATGATGAT 2827
QY 254 cccctcctcgtgactcgtatccgaaatcactccactccactcctcgtcgaactcgaatcga 313
DB 2828 CTAGCTTTTGGATTAAGAAATTTCTAGCTATTTTAACTATTAAGCTTAATTAATTAATTA 2887
QY 314 ccaatcaactcgtcgaagaacacatcctcgttgaagaatcctcgaactcgaactcga 373
DB 2888 CAAATTAATTAATTTATGAGAAATTAATTTAGAGATGAGAAATGATCATTATTAATGAGTAA 2947
QY 374 tcaatcgtgactcgtcgaagaactcgaagaacacacacacacacacacacacacacacac 433
DB 2948 TAATCTGAGCTTTTACAGAGATCTCAGAGATTAATTAATTAATTAATTAATTAATTAAT 3007
QY 434 agatgtaacacatcgtcgtacatcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 493
DB 3008 AAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3067
QY 494 gtcgataactcccaatccatcaatccacacacacacacacacacacacacacacacac 553
DB 3068 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3127
QY 554 atctggttaacatccacacacacacacacacacacacacacacacacacacacacacac 613
DB 3128 ATTTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3187
QY 614 ctcaacccatcactcgtatcaatcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 673
DB 3188 CACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3247
QY 674 aaatcaagaactcgtcgaagaacacacacacacacacacacacacacacacacacacacac 733
DB 3248 AAATCAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3307
QY 734 actaactcgaatcgaagaacacacacacacacacacacacacacacacacacacacacac 793
DB 3308 ATTAATTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3367
QY 794 ttaacgtcaacaatcgtatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 853
DB 3368 TCGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3427
QY 854 ttaagtaacacacacacacacacacacacacacacacacacacacacacacacacacacac 913
DB 3428 TAATGACTCAAAACATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3487
QY 914 agaaatacgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 973

Db	3488	AAAATATGCTTCGGAAATAAAGATTAATTTGTAAGAATAATGATCGTATATATTA	3547
QY	974	atctgttagttaagaacaagaataccgcctcgtgtaccaaatgctctcaagctgtgtag	1033
Db	3548	ATGTAGTAGTGTAAAAATAAAGAAATTAATGGTTAGTACTAATCATCATACAGGCGAGCTAG	3607
QY	1034	aaaagatctgtcctcgtcgtgaaatcccgagcttgtagatcctgcctgaagtagtttaa	1093
Db	3608	AAAAAATACTAAGTCGTTTGAAATATCTGATGTAGGAATAATTAAGTCAAGTAGTAGTAA	3667
QY	1094	tgaatccaaagacagaccagtgatcactaacaatgycanaatgaaatctgcagagcaaca	1153
Db	3668	TGAAGTCAAAAAATGATCAAGGAGTAATTAACAAATTAAGCAAAATGAAATTACAAAGATATA	3727
QY	1154	atgttaacgatatcggtttatcattcgtttccacacagttcaacatatcgctaaactgtgt	1213
Db	3728	ATGGGAATGATTAATAGGCTTTATATAGGAATTCATCATGTTTAATATATAGCTAAACATAGAG	3787
QY	1214	cttccaacttgcatacaactcgcagatcgcgaacagttccctctgcagactctgggttgcctgtg	1273
Db	3788	CAATTAATTTGGTATTAATAGACAATAAGAAAGATCTAGTAGGACTTTGGGTTTCATGGG	3847
QY	1274	agttcatcccggttgtagaagcgtttgggtgtgaacgctccgctgttaa	1317
Db	3848	AATTAATTCCTGTAGATAGATGATGGGGGAAAGGCGCACTGTGTA	3891

RESULT 8
 US-08-405-496A-27
 Sequence 27, Application US/08405496A
 Patent No. 591965
 GENERAL INFORMATION:
 APPLICANT: WILLIAMS, JAMES A.
 TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
 TITLE OF INVENTION: NEUROTOXIN
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 MONTGOMERY STREET, SUITE 2200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/405,496A
 FILING DATE: 16-MAR-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791
 FILING DATE: 31-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: INGOLTA, DIANE E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPBD-01308
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 27:

```

?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 3891 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: double
?      TOPOLOGY: linear
?      MOLECULE TYPE: DNA (genomic)
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 1..3888
?      US-08-405-496A-27

```

[illegible]


```

Db 589 GGTGTGGGGCGCTATTCGGTAGAGCAACAACATCCTTAAAGCTGGAGCCGTTGCAACAC 648
QY 613 actcaccgtacatctcgtgatacaataactcaatctcgtcagacaagaactgaacgaataa 672
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 AACCAACAGTACGTATTCACGACAGATTCGATCTTCTGCAACACACGAAACCCGAA 708
QY 673 gaaataaagacctcgtcagacaaccagtcaccaattctcgtatctcgtgaagaactcgtgg 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 GAGATCGAAAACTGTATTCAGACGTACCTGTATTCACCTTCTCGGTGACCTCTGGGGGT 768
QY 733 gactaccctcagtagacaacacgctactacatcgtgaactcgtgaatcgtcgaacaaatc 792
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 769 AACCCGCTGCTTACGACACCGAATATTACCTGATCCCGTACCTTCTAGCTTAAAGAC 828
QY 793 gttacgtcaacaatgtaagtalccgcggtttacatgtaactgaagaagtcgcggtgttct 852
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 829 GTTCACCTGAAAA-----CATCAGTACTACATGATGTACTGACCAACGGCGGTCTAC 882
QY 853 gttatgactacaacatcttaactgaactcttccctgtaccggtgtacccaatcaccatc 912
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 ACTAACGGTAACCTGAACATCTACTACCGACGTCTGTACAAACGGCTGAAATTCATCATC 942
QY 913 aagaataacgcgtctcgtgtaacaagacaat 942
Db 943 AACGCTACACTCCGAACACGAATTCGAT 972

```

```

RESULT 11
US-08-280-228-3
; Sequence 3, Application US/08280228
; Patent No. 5571694
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHIVE P.C.
; STREET: 1100 NO. 5571694th Glebe Road
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,228
; FILING DATE: 25-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/618,312
; FILING DATE: 27-NOV-1990
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mary J
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 3:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Clostridium tetani
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1356
; US-08-280-228-3

Query Match      12.6%; Score 166.4; DB 1; Length 1359;
Best Local Similarity 52.3%; Pred. No. 3, le-39;
Matches 486; Conservative 0; Mismatches 411; Indels 33; Gaps 4;

QY 40 aacatcatcaatacctccatccctgaacactcgtcgtacgaatccaatcaccatgcagactg 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49 ATCCTGAAAAAGCTAACATTTGAACTTGAGATCATCAACAAGATATTATCTCCGACATC 108
QY 100 tctcgtaagctcccaaaatcaacatccggtctcaagttaactgatacgcg----atcgac 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TCTGTTTCAACTCCTCTGTTATCAGATATTCAGATGCTCAATTGGTCCGGGCATCAAC 168
QY 157 aagaatcagatccagcgtgtcgaatcctcgaatcctcgaatcgaatcctcgaatcgaat 216
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 GGCNAAGCTATCCACCTGGTTAACACAGATCTTGTGAAGTTATCTGTCACAAAGGCCATG 228
QY 217 gcatcgtatacaactctatagcagaatactcaccctcctcgtgatacgtatccg 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 GACATGGAATACAAACGACATGTTCAACAACACTCACCGTTAGCTTGCGGTGCGCGTTCCG 288
QY 277 aaacttaactcaatccatc-----tctcgtgaacaatgaatacaccatcaatcgtc 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 AAAGTTTCGCTGCCACCTGGAACAGTACGGCAGTACAGTACAGTACATCATCATCTCT 348
QY 328 atgaaaaacaat-----tctggttgaaagtatctcgtgaactcagtgtaa 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 ATGAAGAAACACTCCCTGTCATCCGCTGTGGTTGGTCTGTTCCCTGAAAGGTAACAC 408
QY 373 atcatctggaactctcagcagcactcaggaatacaaacagcgtgtgtatcaataactct 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 CTGATCTGGACCTGTGAAGAAGCTCCGGGGAGATTGTCAGATCACTTCCCGACACTG 468
QY 433 cagatgatcaaatctcgtactcaatcaatcgcgtgatactcgttaccatcaccacaat 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 CCGGACAAAGTTCAACGGGTACCTGGGCTTAAACAAATGGGTTTTCATCACTATCACTACGAT 528
QY 493 cgtctgaataacctcaaaatactatacaagcgcgtcgtcagcagaacacgacatcc 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 CGTCTGCTTCTCTTAACCTGTTACATCAACGGGCTTGTATGGCTCCGCTGAATACAT 588
QY 553 aatcgtgtaacatccacgcttctaataacatcatgttcaaacgtgagcgtgtgtcgtac 612
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 589 GGTCTGGGGCTATTCGCTGAGACAAACAACATCACTTAAAGCTGGACCTTGCAACAC 648
QY 613 actcaccgtacatctcgtgatacaataactcaatctcgtcagacaagaactgaacgaataa 672
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 AACCAACAGTACGTATTCACGACAGATTCGATCTTCTGCAACACACGAAACCCGAA 708
QY 673 gaaataaagacctcgtcagacaaccagtcaccaattctcgtatctcgtgaagaactcgtgg 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 GAGATCGAAAACTGTATTCAGACGTACCTGTATTCACCTTCTCGGTGACCTCTGGGGGT 768
QY 733 gactaccctcagtagacaacacgctactacatcgtgaactcgtgaatcgtcgaacaaatc 792
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 769 AACCCGCTGCTTACGACACCGAATATTACCTGATCCCGTACCTTCTAGCTTAAAGAC 828
QY 793 gttacgtcaacaatgtaagtalccgcggtttacatgtaactgaagaagtcgcggtgttct 852

```

Db 829 GTTACAGCTGAAAA-----CATCAGCTACATGATACGACCAACGCCGCCCTCTAC 882
Qy 853 gtatgactaccacatctaccctgaactcttcctcgtgacggtgtaaccaatcatc 912
Db 883 ACTAAGGTAACTAAGTAACTACTACTACGACGCTGTGTACAAAGGCTTAATTCATCATC 942
Qy 913 aagaatacgcgtctgtaacaaggacat 942
Db 943 AAAGCTACACTCCGACACAGCAATGCAT 972

RESULT 12
US-07-618-312A-1
; Sequence 1, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618, 312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1359 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Clostridium tetani
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1356
; US-07-618-312A-1

Query Match 5.6%; Score 74.2; DB 1; Length 1359;
Best Local Similarity 47.38; Pctd. No. 3.9e-12;
Matches 223; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

Qy 340 tctggtgaaagatctctctgacacagtggaatcatctgactctgacgacacacag 399
Db 376 TCTGTTGGAGCTGATCTACTTAAGTAACTTAATTAATGACTTTAAAGATTCCGG 435
Qy 400 gaaatcaaacagcgtgtgtatcaatctcagatgatacaatctctgactacac 459
Db 436 GGAGAAAGTTAGACAATACTTTAGGATTTACCTGATTAATTAATGCTTATTAGCA 495
Qy 460 aatcgtgatactctgtaacatcaacacacatcgtctgaaatcaatccaaatc 519
Db 496 AATTAATGGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 555
Qy 520 aacgacctgctgacacgaacacgactcctcaatcgtgtaacacacacgcttcaat 579
Db 556 AATGAGTACTTATGGAAGTGCAGAAATTTACTGTTTAGGAGCTTATGAGAGTAA 615
Qy 580 aacatcatgttcaaacctgacggtgtcgtgacacacacgctacatctgtaacaa 639
Db 616 AATATACTATTAATACTAGATGATGATTAATTAATTAATTAATTAATTAATTA 675
Qy 640 tcaatcgttgcgaagaactgacgaagaagaagaatcaagacgtgacgaacag 699
Db 676 TTTAGGATATTTTGCAGACATTAATCCAAAGAGATTGAATAATTAATTAATTA 735
Qy 700 tccaatctgatactcgtgaagaactctggtgtagactcgtgacgacgaacacg 759
Db 736 TTAATCTATTAATCTTTTAAGAGACTTCTGGGAACCTTTACGATATATACAGAT 795
Qy 760 taatcgtgacatctgtaacacacacacacacacacacacacacacacacac 810
Db 796 TATTTAATACAGTAGCTTCTAGTTCTAAGATGATGCAATGAAATAATTA 846

RESULT 13
US-08-110-786A-7
; Sequence 7, Application US/08110786A
; Patent No. 5443966
; GENERAL INFORMATION:
; APPLICANT: FAIRWEATHER, Neil Fraser
; APPLICANT: MAKOFF, Andrew Joseph
; TITLE OF INVENTION: Expression of tetanus toxin fragment C.
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 5443966th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110, 786A
; FILING DATE: 23-AUG-1993 1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777, 337
; FILING DATE: 29-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00943
; FILING DATE: 20-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8914122.0
; FILING DATE: 20 June 1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary J. Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100
 TELE: 200797 NIXN UR
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1359 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 ORIGINAL SOURCE:
 ORGANISM: Clostridium tetani
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1359
 US-08-110-786A-7

Query Match 5.6%; Score 74.2; DB 1; Length 1359;
 Best Local Similarity 47.3%; Pred. No. 3.9e-12;
 Matches 223; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 340 tcctgttggaagatctctgaactcaggtgaatcatctgacctcgcaggaactcag 399
 |||||
 DB 376 TCTGTGGAGGTATCACTTAAGTATTAATGACCTTAAAGATTCCGGC 435
 |||||
 QY 400 gaatacaacagcggtgtgtatccaatactcagatgaacaactctgactacatc 459
 |||||
 DB 436 GGAGAGATTAGCAAAATACTTTAGGATTTACCTGATTAATTTAATGCTTATTAGCA 495
 |||||
 QY 460 aatcgctgatactctgctaccacaacaaatcgctgataactcgaacaaatctatc 519
 |||||
 DB 496 AATTAATGGGTTTTTATTAATCTATTAATGATTAATGATTTCTTCTAATTTGATATA 555
 |||||
 QY 520 aacggccgctgctgacgaacgaacgaacatctcaatctggtgaacacgccttaac 579
 |||||
 DB 556 AATGGAGTACTTATGGAGATGACAGAAATTAATGATTTAGAGCTTTAGAGATTAAT 615
 |||||
 QY 580 aacatcatgttcaaatgacggtgtgtgtgacactcaccgcgtacactctgacataac 639
 |||||
 DB 616 AATTAATACCTTTAAACTAGATGATGATTAATTAATTAATCAATGCTTCTATTTGATATA 675
 |||||
 QY 640 tccaatctgtcgaacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaac 699
 |||||
 DB 676 TTTAGATATTTTGGCAAAACATTAATTCMAAAGATTTGAAATTTATACACATTTAT 735
 |||||
 QY 700 tccaatctgtatccctgaacgaacgaacgaacgaacgaacgaacgaacgaacgaac 759
 |||||
 DB 736 TTAATTAATACCTTTTAAAGATCTTCTGGGAAACCTTTAGATATGATACAGAAATAT 795
 |||||
 QY 760 tacaatgtgaatctgtacgacgaacgaacgaacgaacgaacgaacgaacgaacgaac 810
 |||||
 DB 796 TATTTAATACAGTAGCTTCTAGTTCTTAAGATGTTCAATTTGAAAAATATA 846
 |||||

RESULT 14
 US-08-280-228-1
 Sequence 1, Application US/08280228
 Patent No. 5571694
 GENERAL INFORMATION:
 APPLICANT: Makoff Dr, Andrew J
 APPLICANT: Romano Dr, Michael A
 APPLICANT: Clare Dr, Jeffrey J
 APPLICANT: Fairweather Dr, Neil F
 TITLE OF INVENTION: VACCINES
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHYE P C
 STREET: 1100 NO. 5571694th Glebe Road
 CITY: Arlington,
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/280,228
 FILING DATE: 25-JUL-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/618,312
 FILING DATE: 27-NOV-1990
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 8926832.0
 FILING DATE: 28-NOV-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9006097.1
 FILING DATE: 17-MAR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Wilson, Mary J.
 REGISTRATION NUMBER: 32,955
 REFERENCE/DOCKET NUMBER: 117-163
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1359 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Clostridium tetani
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1356
 US-08-280-228-1

Query Match 5.6%; Score 74.2; DB 1; Length 1359;
 Best Local Similarity 47.3%; Pred. No. 3.9e-12;
 Matches 223; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 340 tcctgttggaagatctctgaactcaggtgaatcatctgacctcgcaggaactcag 399
 |||||
 DB 376 TCTGTGGAGGTATCACTTAAGTATTAATGACCTTAAAGATTCCGGC 435
 |||||
 QY 400 gaatacaacagcggtgtgtatccaatactcagatgaacaactctgactacatc 459
 |||||
 DB 436 GGAGAGATTAGCAAAATACTTTAGGATTTACCTGATTAATTTAATGCTTATTAGCA 495
 |||||
 QY 460 aatcgctgatactctgctaccacaacaaatcgctgataactcgaacaaatctatc 519
 |||||
 DB 496 AATTAATGGGTTTTTATTAATCTATTAATGATTAATGATTTCTTCTAATTTGATATA 555
 |||||
 QY 520 aacggccgctgctgacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaac 579
 |||||
 DB 556 AATGGAGTACTTATGGAGATGACAGAAATTAATGATTTAGAGCTTTAGAGATTAAT 615
 |||||
 QY 580 aacatcatgttcaaatgacggtgtgtgtgacactcaccgcgtacactctgacataac 639
 |||||
 DB 616 AATTAATACCTTTTAAACTAGATGATGATTAATTAATTAATCAATGCTTCTATTTGATATA 675
 |||||
 QY 640 tccaatctgtcgaacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaac 699
 |||||
 DB 676 TTTAGATATTTTGGCAAAACATTAATTCMAAAGATTTGAAATTTATACACATTTAT 735
 |||||
 QY 700 tccaatctgtatccctgaacgaacgaacgaacgaacgaacgaacgaacgaacgaac 759
 |||||
 DB 736 TTAATTAATACCTTTTAAAGATCTTCTGGGAAACCTTTAGATATGATACAGAAATAT 795
 |||||

.....

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:30:08 ; Search time 7254.67 Seconds
(without alignments)
3824.928 Million cell updates/sec

Title: US-09-611-419a-5
Perfect score: 1326
Sequence: 1 gattcgaacatgagctc.....aacgtcgcgtgaagattc 1326

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_pa:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
------------	-------------	-------	--------	----	----	-------------

1	1302.2	98.2	1330	6	AR000029	AR000029 Sequence
2	1302.2	98.2	1330	6	AR169140	AR169140 Sequence
3	1302.2	98.2	1330	6	AX036243	AX036243 Sequence
4	1302.2	98.2	1338	12	XXU22962	022962 Synthetic b
5	1302.2	98.2	1402	6	AR000030	AR000030 Sequence
6	1302.2	98.2	1402	6	AR169141	AR169141 Sequence
7	1302.2	98.2	1402	6	AX036246	AX036246 Sequence
8	1302.2	98.2	1402	6	CB007A	X52066 Clostridium
9	727.2	54.8	4835	1	CLONEUR	M30196 C.botulinum
10	726	54.8	3891	6	AR000031	AR000031 Sequence
11	726	54.8	3891	6	AR169142	AR169142 Sequence
12	726	54.8	3891	6	AX036248	AX036248 Sequence
13	661.6	49.9	4067	1	CBNT0XA	X73423 C.botulinum
14	423.6	31.9	1299	12	AF251281	AF251281 Synthetic
15	422	31.8	1313	6	A58946	A58946 Sequence 6
16	227.8	17.2	4073	1	CBONTF	X68262 C.barati ge
17	227	17.1	1293	6	A58945	A58945 Sequence 5
18	227	17.1	4199	1	CLOBONT	L35496 Clostridium
19	227	17.1	4209	1	CBONTFG	X81714 C.botulinum
20	194.8	14.7	3835	1	AB037704	AB037704 Clostridi
21	194.8	14.7	3835	1	AB037705	AB037705 Clostridi
22	194.8	14.7	3835	1	AB037706	AB037706 Clostridi
23	194.8	14.7	3835	1	AB037707	AB037707 Clostridi
24	194.8	14.7	3835	1	AB037708	AB037708 Clostridi
25	194.8	14.7	3835	1	AB037709	AB037709 Clostridi
26	194.8	14.7	3835	1	AB037710	AB037710 Clostridi
27	194.8	14.7	3835	1	AB037711	AB037711 Clostridi
28	194.8	14.7	3835	1	AB037712	AB037712 Clostridi
29	194.8	14.7	3835	1	AB037713	AB037713 Clostridi
30	194.8	14.7	4030	1	AB037714	X62683 C.botulinum
31	194.8	14.7	4030	1	CBNTTE	X62089 C.botulinum
32	193.2	14.6	4017	6	AX088262	AX088262 Sequence
33	193.2	14.6	3836	1	AB039264	AB039264 Clostridi
34	190	14.3	3949	1	CBNTTE	X62088 C.botulinum
35	190	14.3	4234	1	CLONEUOXF	M92906 Clostridium
36	180.8	13.6	9325	1	CBY13631	Y13631 Clostridium
37	180.8	13.6	9325	6	I28431	I28431 Sequence 3
38	166.4	12.5	1359	6	A49987	A49987 Sequence 4
39	166.4	12.5	3712	6	A49987	A37074 Sequence 17
40	166.4	12.5	3754	6	A37074	A42478 Sequence 17
41	166.4	12.5	3754	6	A42478	A37075 Sequence 18
42	166.4	12.5	3769	6	A37075	A49988 Sequence 5
43	166.4	12.5	4366	6	A49988	A42484 Sequence 12
44	166.4	12.5	4366	6	A42484	A42481 Sequence 9
45	166.4	12.5	4378	6	A42481	

ALIGNMENTS

RESULT 1
LOCUS AR000029 1330 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5736139.
ACCESSION AR000029
VERSION AR000029.1 GI:3962560
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1330)
Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.
TITLE Treatment of Clostridium difficile induced disease
JOURNAL Patent: US 5736139-A 22 07-APR-1998;
FEATURES
source Location/Qualifiers
1..1330
/organism="unknown"
BASE COUNT 400 a 339 c 246 g 345 t
ORIGIN

Query Match 98.2%; Score 1302.2; DB 6; Length 1330;
Best Local Similarity 99.8%; Pred. No. 0;

	Matches	1304:	Conservative	0:	Mismatches	3:	Indels	0:	Gaps	0
QY	14	tgcccttcaaccttcaactgaataatacatcaagaataatcatataatccctcccatccctgaacctgc	73							
Db	11	tgctcttcaaccttcaactgaataatcatcaagaamaatcatcaatctccatccctgaacctgc	70							
QY	74	gtactgaatccaatacaacctgatacgcacctgctctgcgtactacgcttccaaataatcaacatcoggt	133							
Db	71	gcttagaattccaatcaacctgaatcgacacctgtctgcgcttagccttccaaaatcaacatcoggt	130							
QY	134	ctaaagttaacctcgatcccgatcgacaagaatcagatcccaagctgtlccaatctgtgaatcct	193							
Db	131	ctaaagtttaacctccatccgatccgatccgaagaatcaatccatccagctgttcaatctggaatctt	190							
QY	194	ccaaaatcgaaagttaacctgaagaatgtatcgtatacaacctctgtagcgaanaacttct	253							
Db	191	ccaaaatcgaaagttaacctgaagaatgtatcgtatgatacaacctctgtatgcaaaaacttct	250							
QY	254	ccaacctctctcgtatcccgatcccgaaataacttaactccaatctctctcgtacaaatgaat	313							
Db	251	ccacctctctcgtatcccgatcccgaaataacttaactccaatctctctcgtacaaatgaat	310							
QY	314	aacacatatacaactgcgaatggaacaaatttcggtltgaaagtaactctgaaactaacgtg	373							
Db	311	aacacatatacaactgcgaatggaacaaatttcggtltgaaagtaactctgaaactaacgtg	370							
QY	374	aaatcatctcgtaccttgaagaagacacacacacgaagaatacaaacagctgtgtgtatccaatct	433							
Db	371	aaatcatctcgtaccttgaagaagacacacacacgaagaatacaaacagctgtgtgtatccaatct	430							
QY	434	ctcgaatgatacaaatctctgactataccaatcgcgttgatctcgtttaccataccaacaa	493							
Db	431	ctcgaatgatacaaatctctgactataccaatcgcgttgatctcgtttaccataccaacaa	490							
QY	494	atcgcctgaataaactcccaaaatactacatcaacagcgctctga tgcacgaagaaccgatct	553							
Db	491	atcgcctgaataaactcccaaaatactacatcaacagcgctctga tgcacgaagaaccgatct	550							
QY	554	ccaatctcgtgttaacatccaagctcttcaataccaatcgaatctgaactcgtgacgtgtg	613							
Db	551	ccaatctcgtgtgttaacatccaagctcttcaataccaatcgaatctgaactcgtgacgtgtg	610							
QY	614	acaactcaacgcgtatactctgatacaataactccaatctcttgcgaanaaagacgaaacgaa	673							
Db	611	acactcaacgcgtatactctgatacaataactccaatctcttgcgaanaaagacgaaacgaa	670							
QY	674	aagaataccaagaacacctgtacgacaacaacagctccaatctctgtlatacctgaagaacctctg	733							
Db	671	aagaataccaagaacacctgtacgacaacaacagctccaatctctgtlatacctgaagaacctctg	730							
QY	734	gtgactacctcgtcagtagaagaacacccgtactatactgatactctgtacgatccgaacaaat	793							
Db	731	gtgactacctcgtcagtagaagaacacccgtactatactgatactctgtacgatccgaacaaat	790							
QY	794	acgltcgaagcacaacatgtlaagatcgcgcggttatacatgttacctgaaagaagctccgcgtggt	853							
Db	791	acgltcgaagcacaacatgtlaagatcgcgcggttatacatgttacctgaaagaagctccgcgtggt	850							
QY	854	ctgttatgtactacaacaactctacactgaactcttccctgtacgcgtgtgtacaaatcatca	913							
Db	851	ctgttatgtactacaacaactctacactgaactcttccctgtacgcgtgtgtacaaatcatca	910							
QY	914	tcaagaanaatacgcgtctgtlataacaagaagacaatacgttctgcgaacaatgtatgtlataca	973							
Db	911	tcaagaanaatacgcgtctgtlataacaagaagacaatacgttctgcgaacaatgtatgtlataca	970							
QY	974	tcaatgtgtgtatgttaagaacaagaataacacgtctgtgtctaccaatgtcttcacgctcgtg	1033							
Db	971	tcaatgtgtgtatgttaagaacaagaagaatacacgctgtgtctaccaatgtcttcacgctcgtg	1030							
QY	1034	tagaanaaatacttctgtctctctcgtgaatacccggaacggttgtgtaactctgtctcgaagctgtg	1093							
Db	1031	tagaanaaatacttctgtctctctcgtgaatacccggaacggttgtgtaactctgtctcgaagctgtg	1090							

Query Match	Best Local Similarity	Matches 1304; Conservative	98.2%; Score 1302.2; DB 6; Length 1330; Pred. No. 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1094	14	14	14
Db 1091	11	11	11
Qy 1154	74	74	74
Db 1151	71	71	71
Qy 1214	134	134	134
Db 1211	131	131	131
Qy 1274	191	191	191
Db 1271	254	254	254
Qy 1314	311	311	311
Db 1311	374	374	374
Qy 1371	434	434	434


```

Oy 674 aagaatacaagacgtgtacagcaaccagtcacattctgtatccttgaagaactctggg 733
    |||
Db 671 AAGAAATCAAAAGACGTGTACGACCAACGACCAATTCGTGTATCTGAAAGACTTCTGGG 730
Oy 734 gttactactcgtgtagtgcgaacacgttacttaactgttgaattctgttcggtccgaacaat 793
    |||
Db 731 gttactactcgtgtagtgcgaacacgttacttacttgcgtgaattctgttcggtccgaacaat 790
Oy 794 acgttgcgttcaacaactgtatcgtatccgcgtgtatcactgtacacgttcacgttcgttcgt 853
    |||
Db 791 AGTTGACGTCAACATGTAGTATCGGGTTACATGTACCTGGAAGGTCGCCGTGTT 850
Oy 854 cgttatgactcaccacattactactcgttaactcttcctgtacacgttggtaaccaatcata 913
    |||
Db 851 CTTGATGTACTACCAACATCTACTACTGACTTCCCTGTACCTGCTGACCAAAATTCATCA 910
Oy 914 tcaagaatacagcgtctgttgaagaagaatactgttcgcaacaatgtatcgtatata 973
    |||
Db 911 TCAAGAAATACGGCTTGTGTACAGGACAAATATCGTTGCAACAAATGATCGTATACA 970
Oy 974 tcaatgttgaatlaagaacaagaatacagtcgttcacacgttccttcaagcgtgtg 1033
    |||
Db 971 TCAATGTGTAGTATAGAACAAAGATACGCTGTGCTACCAATGCTTCTCAGGCTGTT 1030
Y 1034 tgaagaagatctgtctgtctcgtcgtgaatacccggaagttgtgtaactgtctcaagtgtg 1093
    |||
Db 1031 TAGAAAAGATCTTGTCTGTCTGGAATCCCGACCTTGTGTATCTGCTCAGGTAGTTG 1090
Oy 1094 taatgaatacacaagaacagcaggtgtatcacttaacaataatgaatctgcagaaca 1153
    |||
Db 1091 TATATGAATCAAGAACGACCAAGGTATCTATCAATGCAAAATGCAAAATCTCAGAGACA 1150
Oy 1154 acaatgttaacgatatcgttctcgttcgttcacacagttccacaatalatcgtctaaactg 1213
    |||
Db 1151 ACAATGTGTACGATATCGGTTTCATCGGTTTCACACAGTTCAACAAATGCTGTAACCTG 1210
Oy 1214 ttgcttccacacgtgtacatcgttcaatcgaatcgttcctcgtcgtcgtgtgtctct 1273
    |||
Db 1211 TTGCTTCCACACTGTGTACAAATCGTCAATGTGAACGTTCTCTGCACTGTGGTGTCTTT 1270
Oy 1274 gggagttacatcccggtgtgtatgtagcgtgtggtgtagaacgttcgctgtata 1320
    |||
Db 1271 GGGAGTTATCCCGGTTGTGATGACGGTTGGGTTGAACGTCCGCTGTAA 1317

RESULT 4
LOCUS XXU22962 1338 bp DNA linear SYN 08-NOV-1995
DEFINITION Synthetic botulinum neurotoxin serotype A Hc fragment (botA) gene,
complete cds.
ACCESSION U22962.1 GI:733428
VERSION 1
KEYWORDS
SOURCE
ORGANISM
.
synthetic construct.
synthetic construct
artificial sequence.
1 (bases 1 to 1338)
REFERENCE
AUTHORS Clayton, M.A., Clayton, J.M., Brown, D.R. and Middlebrook, J.L.
TITLE Protective vaccination with a recombinant fragment of Clostridium
botulinum neurotoxin serotype A expressed from a synthetic gene in
Escherichia coli
JOURNAL Infect. Immun. 63 (?), 2738-2742 (1995)
MEDLINE 95310035
REFERENCE 2 (bases 1 to 1338)
AUTHORS Brown, D.R.
TITLE Direct Submision
JOURNAL Submitted (17-MAR-1995) Douglas R. Brown, Toxicology, U. S. Army
Medical Research Institute of Infectious Diseases, Bldg. 1425, Fort
Detrick, Frederick, MD 21702-5011, USA
LOCATION/Qualifiers
FEATURES
source
1..1338
/organism="synthetic construct"
/db_xref="taxon:32630"

```

```

/lab_host="Escherichia coli"
/note="based on Clostridium botulinum Type A neurotoxin
sequence"
gene
9..1325
/gene="botA"
9..1325
/gene="botA"
/codon_start=1
/transl_table=11
/product="botulinum neurotoxin serotype A Hc fragment"
/protein_id="AAA80610.1"
/db_xref="GI:733429"
/translation="MARLSTFEYIKNIINTSINLEYESNHLIDLSRYAKINIGS
KVPDPIDKNQIOLFNESSKLEVLKNAIVNSYENSTFSFIRIPKYSISLN
EYTIINCMENSGMYSLNGEIIWTLQTOEIKORVYKROMINISDY IIRWLEVT
ITNNRLNSKIIYINGRLIDOKPISNGINHASNNIMFKLDGCRDHRHYIRNPLFD
KELNKEIKDIYDNDOSNGSLIKDPNGDYLOKPYVMNLXPDKYVYVNNYJEGYM
YIKGPRGSVMTNTNITUNSLYKGTFTIKKIASGKNIDIVRNDRYIVNKKRYR
LATNNSQAVEKILSALEIPDVGNLSQVVVMSKNDQGITNCKMNLQNDNGDIGFI
GPHQNNIAKLIVASWYNNROIERSRSRTLCSSWEFLPVDGNGERPL"
BASE COUNT 401 a 342 c 249 g 346 t
ORIGIN

```

```

Query Match 98.2%; Score 1302.2; DB 12; Length 1338;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Oy 14 ttgcccctaccctcactgtagatatacaagaactatcaactcactcgtgaactgc 73
    |||
Db 19 TGTGTCTACTCTTACCTGATGATACATCAAGAACATCATCAATACCTCCTCAATCCATCG 78
Oy 74 gctacgaatcaaatcacctgatcgactgtctcgtacagcttccaaatatcaaatcgtgt 133
    |||
Db 79 GCTACGAATCAATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138
Oy 134 ctaaatlaacttcgatccgatcgtacagaagaatcgaatcgaatcgaatcgaatcgt 193
    |||
Db 139 CTTAAATCTTAATCTTGAATCCGATCGACAGAAATGATGATGATGATGATGATGATGATGAT 198
Oy 194 ccaaatcgaagtattccctgaagaatgctatgatacaactatgtaagaagaacttct 253
    |||
Db 199 CCAAAATCGAAGTTATCTTGAAGAAATGCTATGATGATGATGATGATGATGATGATGATGAT 258
Oy 254 ccaactcctcttgatccgtatccgaaatcactcaactcactcactcactcactcactcact 313
    |||
Db 259 CCACCTCTCTTGATGATCGGATCCGAAATATCTCAACTGATGATGATGATGATGATGATGAT 318
Oy 314 acacatcatcaactcgtcatggaagaactcgtgtggaagaatgatatctcgtgaactcgt 373
    |||
Db 319 ACACATCATCAATCACTGATGGAAGAAATCTGTTGGAAGATATCTGAACTGATGATGATGAT 378
Oy 374 aaatcatcgtgactcgtcaggaacactcagaagaatcagaagcgtgtgtatataact 433
    |||
Db 379 AAATCATCTGACTCTGCAAGGACCTCAGGAATTCAGGACGCTGTGTTGTTGTTGTTGTTGTT 438
Oy 434 ctcaagaatcaaatcactcgtatcaatcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 493
    |||
Db 439 CTCAGATGATCAACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 498
Oy 494 atcgtctgaataactccaaatctacatacagaagcgtcgtcgtcgtcgtcgtcgtcgtcgt 553
    |||
Db 499 ATCGTCTGAATTAATCTCAAAATCTATCATCAACGCGCTGTGATGACACGAAACCGATCT 558
Oy 554 ccaatcgtgttaacatccacacgtcttcaataactatgttcaaacctgagcgtgtgtcgtg 613
    |||
Db 559 CCAATCTGGTATCAATCCAGGCTTCTTAATTAATCAATGTTCAACCTGAGGAGGTTGTCGTG 618
Oy 614 acactacgcgtacactcgtatcaatcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 673
    |||
Db 619 ACACCTACCGGTTACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
Oy 674 aagaatacaagacgtgtacgacaaccagtcacattctgtatccttgaagaactcgtgg 733

```



```

|||||
Db 679 AAGAAATCAAGACCTGTACGACAAACCAATCTGATTCGGAAGACTTGTGG 738
Oy 734 gtagactactgtagtagcaagaacogtatactatgctgaatctgtacgctccgaacaa 793
Db 739 GTGACTACCTGGAGTACGCAAAACCGTACTACATCTGTATCTGTACGATCGGAACAAAT 798
Oy 794 acgttagacgtaacaatgtagtaltccgcggtgtacatgaactgaaggtccgcgtggt 853
Db 799 ACCTTGACGTCAACAATGATAGTATCCGGTTTACATGACTGAAAGGTCCGGCTGTT 858
Oy 854 ctgttatgaactacaacatctacccctgaactctccctgtacgctgtgtacaaatcata 913
Db 859 CTGTATGACTACCAACATCTACCTGAATCTCCCTGTACCGTGTACCAATTCATCA 918
Oy 914 tcaagaataacgctctgtgtacagaagaataatgctgtgcaacaatgctgtgtataa 973
Db 919 TCAGAAATAACCGCTGTGTACCAAGGACAAATGCTTGCAACAATGATCGTGTATACA 978
Oy 974 tcaatgtgtagtagaagaacaagataacgctgtgtacccaatgtcttcaggcgtgtg 1033
Db 979 TCAATGTTTACTTAAAGAACAAAGATACCGTCTGCTACCAATGCTTTCAGGCTGGTG 1038
Oy 1034 tagaaaagaatctgtctgtctgtgaatcccggaacgtgtgtatctgtctcagtagt 1093
Db 1039 TAGAAAAGATCTTGCTGCTGTGAAATCCGAGCTTGTAATCTGTCTCAGGTAGTTG 1098
Oy 1094 taatgaataccaagaagccagggatatacctaacaataatgcaaatgaaatctgcagaca 1153
Db 1099 TAATGAATAATCCAAAGACGACAGGATATCCTACAAATGAATGAAATCTGCAGAGACA 1158
Oy 1154 acaatgtgaagatatacggtttcaccggtttccacaagttaacaatatcgctaaatcg 1213
Db 1159 ACAATGTGATCAATTCGCTTTCATCGGTTTCACAGTTCAACAATTCGCTAAACTGG 1218
Oy 1214 ttgcttccaactgtgacaaatcgtlcaagatcgaacgttctctcgcactcgtgtgtct 1273
Db 1219 TTGCTTCCAACTGTGACATCTGTCAGATGCAAGCTTCTCTGCACTCTGGGTTGCTCT 1278
Oy 1274 gggaaattcaccggtgtgtacggttgggtgtggaacgtccgctgttaa 1320
Db 1279 GGGAGTTCAATCCCGGTTGATGACGGTTGGGGTGAACGTCCGCTGTAA 1325

RESULT 5
AR000030 1402 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 25 from patent US 5736139.
ACCESSION AR000030
VERSION AR000030.1 GI:3962561
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.
TITLE Treatment of Clostridium difficile induced disease
JOURNAL Patent: us 5736139-A 25 07-Apr-1998;
FEATURES
Location/Qualifiers
source
1..1402
/organism="unknown"
BASE COUNT 420 a 360 c 260 g 362 t
ORIGIN

```

Query Match 98.2%; Score 1302.2; DB 6; Length 1402;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Oy 14 tggccttaacttaactatatacaagaacaataataatctccatcctgaactg 73
Db 83 TGCTGTACCTTCACTGATACATCAAGAACATCATCAATACCTCCATCTGAACTTCG 142

```

```

Oy 74 gctacgaatccaatcacactgtagactgtctcgtacgcttccaaatacaatcaggt 133
Db 143 GCTACGAATCCAATCAATCACTGATGACCTGTCTGCTGAGCTTCCAAAATCAATCGGTT 202
Oy 134 ctaagttaaactcgtatccgatcgcataagaataatccagctgttcaatctggaact 193
Db 203 CTAAAGTTAACTTCGATCCGATCGACAAAGATCAGATCCAGGTGTTCAATCTGGAATCT 262
Oy 194 ccaaaatgaagttaatccctgaagaatgctatcgtatacaactctatgtcgaagaact 253
Db 263 CCAAAATCGAAGTTATCTGGAAGAAATGCTATATCAACTCTATGTACGAAACTTCT 322
Oy 254 caacctcctctgtgacgcgtatccggaataactcaaccctcctcctcgtgaacaatgaa 313
Db 323 CCACCTCTTCTGATCCGATTCGCAATFACTTCACTCCATCTCTCTGAAACAAATGAA 382
Oy 314 aacaatcatcaactgtgcatggaanaaacttcgtgtggaaglatcctcgaactacggt 373
Db 383 ACACCATCATCAACTGCATGGAATAAACAAATTCGTGGAAGATCTCTGAACCTACGGTG 442
Oy 374 aatatctggaactctgcaggaacactcaagaaatacaacaggtgtgtatccaatct 433
Db 443 AATCATCTGGACTCTGACAGACACTCAGGAATCAACAGCGTGTGTATTCAAATCT 502
Oy 434 ctcaagatgaatcaatctctgactatcatcatcgtgtgtctcgttaccatcaaca 493
Db 503 CTCAGATGATCAACATCTCTGACTACATCATATCGCTGGATCTTCTGTACATCACCACA 562
Oy 494 atcgtctgaataactccaataatcatatacaacagccgctgtgtacgacgaacagatct 553
Db 563 ATGCTGTGAATTAATCTCCAAATCTACATCAACGGCGCTGTGATCGACAGAAACGATCT 622
Oy 554 ccaatctgtgtatacaatcaacagctctctaataataatgttcaaacctgagcgtgtgt 613
Db 623 CCAATCTGGGTAAACATCCACGCTTCTTAATTAATCAATCAATGTTCAAACTGGACG 682
Oy 614 acaatcagcgtacatctgtatcaaatctcaatctgtctcgaagaagcgaagaa 673
Db 683 ACATCTACCGCTACATCTGTGATCAATTAATCTGTGTCGACAAAGAACTGAAGAA 742
Oy 674 aagaataccaagaactgttgcgaacaacagltccaatctgtgtatcctgtgaagaact 733
Db 743 AAGAAATCAAGACCTGTGACGACAAACCAATCTGATCTGATTCGAAAGACTTGTGG 802
Oy 734 gtagactactgtagtagcaagaacogtatactatgtgtatctgtacgctccgaacaa 793
Db 803 GTGACTACCTGGAGTACGACAAACCGTACTACATGCTGATGTGATCCGAACAAAT 862
Oy 794 acgttagacgtaacaatgtagtagatccgcggtgtatcgtacgtgaagaagtcgcgtgt 853
Db 863 ACCTTGACGTCAACAATGATAGTATCCGGGTTTACATGTACTGAAAGGTCCGGCTGTT 922
Oy 854 ctgttatgaactacaacatctacccctgaactctccctgtacgctgtgtacaaatcata 913
Db 923 CTGTATGACTACCAACATCTACCTGAATCTCCCTGTACCGTGTACCAATTCATCA 982
Oy 914 tcaagaataacgctgtgtgtacagaagaataatgctgtgcaacaatgctgtgtataa 973
Db 983 TCAGAAATAACCGCTGTGTACCAAGGACAAATGCTTGCAACAATGATCGTGTATACA 1042
Oy 974 tcaatgtgtagtagaagaacaagataacgctgtgtacccaatgtcttcaggcgtgtg 1033
Db 1043 TCAATGTTTACTTAAAGAACAAAGATACCGTCTGCTACCAATGCTTTCAGGCTGGTG 1102
Oy 1034 tagaaaagaatctgtctgtctgtgaatcccggaacgtgtgtatctgtctcagtagt 1093
Db 1103 TAGAAAAGATCTTGCTGCTGTGAAATCCGAGCTTGTAATCTGTCTCAGGTAGTTG 1162
Oy 1094 taatgaataccaagaagccagggatatacctaacaataatgcaaatgaaatctgcagaca 1153
Db 1163 TAATGAATAATCCAAAGACGACAGGATATCACTAACAAATGAATGTGCAGAGACA 1222
Oy 1154 acaatgtgaagatatacggtttcaccggtttccacaagttaacaatatcgttaaacgt 1213

```

Db 1223 ACAATGGTAACGATATCGGTTTCATCGTTCCACCAAGTTCAACATATCGCTAACTGG 1282
 Qy 1214 ttgcttccaaactggtacaaatcgctcagatcgaaactcctctgcatctgggttgcctt 1273
 Db 1283 TTCTCTCCAACTGGTACAACTGTCAGATCGCAAGCTTCTCTCGCACTCGGGTGTCTT 1342
 Qy 1274 gggagttcaatcccggttgatgaagcttgggggtgaacgtccgctgtaa 1320
 Db 1343 GGGAGTTCAATCCCGGTTGATGACGGTTGGGCTGACGTCGCTGTAA 1389

RESULT 6
 ARI69141 1402 bp DNA linear PAT 17-DEC-2001
 LOCUS ARI69141
 DEFINITION Sequence 25 from patent US 6290960.
 ACCESSION ARI69141
 VERSION ARI69141.1 GI:17906911
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1402)
 AUTHORS Kink, J. A., Thalley, B. S. and Stafford, D. C.
 TITLE Vaccine and antitoxin for the treatment of C. difficile disease
 JOURNAL Patent: US 6290960-A 25 18-SEP-2001;
 FEATURES
 source 1..1402
 /organism="unknown"
 BASE COUNT 420 a 360 c 260 g 362 t
 ORIGIN

Query Match 98.2%; Score 1302.2; DB 6; Length 1402;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 ttgcttctactcctcaatgaatacaagaacatcaatcaatcctccatctcgtgaactgc 73
 Db 83 TGCTGTCTACCTTCACTGATCATCAAGAACATCATCACTCCATCTCGAACCTGC 142
 Qy 74 gctgaatccaaatcaactcgtatcgactctgctgaagcttccaagaatgaatgcgtt 133
 Db 143 GCTACGAATCCAACTCACTGATGACCTGTCTGCTACGCTTCCAAATCAATCAATCGTT 202
 Qy 134 ctaaagttaactcgtatccgatcgacaagaatcagatccagctgtccaatctgaaactc 193
 Db 203 CTAAAGTTAACTTGCATCCGATCGACAGATCAGATCCAGCTGTTCATCTCGAATCTT 262
 Qy 194 ccaaatcgaaatctatccgtgaagaatgctatcgatatacaactctatgtacgaactctc 253
 Db 263 CCAAAATCGAAATTAATCTGMAAGATGCTATCTATACAACTCTATGTAGCAAAACTTCT 322
 Qy 254 ccaactcctctcgtatccgctatccgaaatcaatcaatcctccatctcgtgaactgaat 313
 Db 323 CCACCTCTCTTCTGGATCCGATCCCGAAATATCTCAACTCTCTGTAACAAATGAAT 382
 Qy 314 acacatcatcaactcgtatgaaacaatctcgttggaaagtatcctgaactcgtg 373
 Db 383 ACACCATCACTCACTGATGGAAGAAACAATCTCTGTTGGAAGATCTCTGAACCTACCGTG 442
 Qy 374 aaatcatctgactctgcaagacactcgaagaatcaacaagcgtgtgtatcaatcaat 433
 Db 443 AAATCATCTGAGACTCTGCAAGACACTCAGGAATCAAAACAGCTGTGTATCAAAATACT 502
 Qy 434 ctaagatgtaacaactctcgtatcaatcaatcgtcgtgactcgttaccatcaacaaca 493
 Db 503 CTGAGATGATCAACATCTCTGACTACATCAATGCTGGAGATCTTCTGACCATACCAACA 562
 Qy 494 atcgtctgaataactccaataatcatcaacagcgcgtctgacgcgaaccagaacgatc 553
 Db 563 ATCGTCTGAATTAATCTCAAAATCTACATCAACGCGCTGTGATGACACGAAACCGATCT 622

Qy 554 ccaatctgggttaacatccacgcctctctaataacatcgttcaaaactggaacggttgcgtg 613
 Db 623 CCAATCTGGGTAACTACATCCAGGCTTTAATTAATCATATGTTCAAACTGAGCGGTTGCTGG 682
 Qy 614 acactcaccgctatcatctggtatcaaatcactcaatctgttgcgaagaactggaagaa 673
 Db 683 ACACCTACCGCTACATCTGATCAAAATTAATCTCAATGTGTCGCAAAAGAACTGAACGAAA 742
 Qy 674 aagaataccaagaacgcttaccgaacaacgctccaatctcgtatccctggaagaactctgg 733
 Db 743 AAGAAATCAAGACGCTGTAGACACACAGTCAATCTGATCTCGTAACGAAAGACTTCTGGG 802
 Qy 734 gtaactaccctgcaatcagacaacacgtaactatcgtcgaatctgtacgatccgaacaa 793
 Db 803 GTGACTACCTCGCATGACGACAAACCGTACATCATGCTGAATCTGTACGATCCCAACAAAT 862
 Qy 794 agcttgaactcaacaatctgaatcgcggttgaactcgtgaactgaagaagctccgctggt 853
 Db 863 ACCTTGAACGTCAACAAATGATAGGTATCCGCGGTACATGATCTGAAAGTCCGCGGTGT 922
 Qy 854 ctgtatgaactcaacaacatcactgaactctccctgtacacgcttggtaaccaatlcata 913
 Db 923 CTGTTATGACTACCAACATCTACTGTAAGTCTCCGTGATACCGTGTACCAATTCATCA 982
 Qy 914 tcaagaataacgctctgtgtacacaaggaacaatcgttgcgaacaatgctgtatata 973
 Db 983 TCAAGAAATACGCGCTGTGTAACAGGACAAATATCGTTCCCAAAATGATCGGTATACA 1042
 Qy 974 tcaatgttgaatgaagaacaagaataacgctctgctgaactcgttccaaatgcttccgaagcgtg 1033
 Db 1043 TCATGTTGTGATTAGAACAAAGAAATACCGTGTGGCTACCAATGCTTCTCAGGCTGGTG 1102
 Qy 1034 tagaagaatctgtctcgtccttgaatcccggaacgcttggtaactcgtctcgaagtagtgc 1093
 Db 1103 TAGAAAAAGATCTTGTCTGCTCGGAATCCCGACGCTGGTATCTGTCTCAGAGTGTG 1162
 Qy 1094 taatgaatccaaagaacgacgaaggtatcactatacaaatgcaaatgtaactcgcagaca 1153
 Db 1163 TAATGAATTCAGAAAGACGACCAAGGTATCACTTAACAAATGCAAAATGCAATCTCAGAGACA 1222
 Qy 1154 acaatgttaacgatactcgtgttcatcggttccacgaagttcaacaatalcgtgaactg 1213
 Db 1223 ACAATGGTAACGATATCGGTTTCATCGGTTCCACCAAGTTCAACATATCGCTAACTGG 1282
 Qy 1214 ttgcttccaaactggtacaaatcgctcagatcgaaactcctctgcatctgggttgcctt 1273
 Db 1283 TTCTCTCCAACTGGTACAACTGTCAGATCGCAAGCTTCTCTCGCACTCGGGTGTCTT 1342
 Qy 1274 gggagttcaatcccggttgatgaagcttgggggtgaacgtccgctgtaa 1320
 Db 1343 GGGAGTTCAATCCCGGTTGATGACGGTTGGGCTGACGTCGCTGTAA 1389

RESULT 7
 AX036246 1402 bp DNA linear PAT 16-NOV-2000
 LOCUS AX036246
 DEFINITION Sequence 25 from Patent EP1041149.
 ACCESSION AX036246
 VERSION AX036246.1 GI:11225864
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 1402)
 AUTHORS Kink, J. A., Firca, J. R., Padhye, N. V., Thalley, B. S., Stafford, D. C. and Williams, J. A.
 TITLE Vaccine and antitoxin for treatment and prevention of C. Difficile disease
 JOURNAL Patent: EP 1041149-A 25 04-OCT-2000;
 FEATURES
 source 1..1402
 /organism="synthetic construct"

Db 2650 GAAAGTATCATTTAATAGACTTATCTAGTATGCATCAAAAATTAATTTGTAGTAAA 2709
 Qy 139 gtttaactcgcgtatcgatcgacaagaatcagatccgcgttcaactcgtgaactctccaa 198
 Db 2710 GTAAATTTTGTATGAAATGATTAATAAATCAATTAATTAATTTAAATTTGTAAGTAA 2769
 Qy 199 atcgaagttacccgtgaagaatgctatcgatatacaactatgtaagaactccccc 258
 Db 2770 ATTGAGTAATTTTAAATAAATGCTATGTATATATATAGTATATGAAAATTTTGTACT 2829
 Qy 259 tccctcgtatccgtaaccgaataactccaactccctcgtgaacaatgatacc 318
 Db 2830 AGCTTTTGATGATGAAATTCCTTAAGATTTTAAACAGTATAGCTTAATATGAATATACA 2889
 Qy 319 atctcaactcgtggaagaacaattcgttgggaagratcctcgtactcgtgaactc 378
 Db 2890 ATAAATTTTGTATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2949
 Qy 379 atctggaactcgtcgaagacacgcgaagaatcaacacgcgtgtgtatcaactcctcag 438
 Db 2950 ATCTGACTTTTACAGATACACGAGAAATTAACAAAGAGTAGTTTTAAATACAGTCAA 3009
 Qy 439 atgatacaactcgtgactacatcaactcgtcgtgactcgttaccataccaactcgt 498
 Db 3010 ATGATTAATATATCATGATTAATTAATAACAGATGATTTTGTAACTATCACTAATATAGA 3069
 Qy 499 ctgaaataactccaatactacatcaacagcgctcgtatcgaccagaaccgactccat 558
 Db 3070 TTAATATATCTTAATAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 3129
 Qy 559 ctgggttaacatccacgccttcaataacatcgttcaactcgtgaagcgttgcgtacact 618
 Db 3130 TTAGGTATATTTACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3189
 Qy 619 caccgctacatcgtgatacaactcgtcgttgcgaagaactcgaagaagaaga 678
 Db 3190 CATGATTAATTTTGTATGATTAATTAATTTTGTATTAATTTTGTATTAATTAATTA 3249
 Qy 679 atcgaagacgtcgtgaagaacacgcgtcgtatcgttgcgaagaactcgtgagctac 738
 Db 3250 ATCAAGATTTTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3309
 Qy 739 tactcgtcgtatcgacaacacgcgttactacatcgtcgtatcgtgaacacgaactcgt 798
 Db 3310 TATTTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3369
 Qy 799 gaactcgaacatgtagtataccgcgttactacatcgtgaagaactcgtgagctac 858
 Db 3370 GATGTAATTAATTTGATGATTAATTAATTAATTTTGTATTAATTTTGTATTAATTA 3429
 Qy 859 atgctacacacatcctactcgaactcctccctcgtacgcgttgcgaagaactcctcga 918
 Db 3430 ATGACTTAACAAATTTTATTTAATTTCAAGTTTGTATGAGGACAAATTTAATTTAATA 3489
 Qy 919 aataacgcgtcgtgaagaacacatacgttgcgaagaactcgtgaactcgtatcgaact 978
 Db 3490 AAAATGCTTCTCGAATAATTAATTAATTAATTTGTTAATAATTAATGATCGTATATTTA 3549
 Qy 979 gttgtagttaaagaacaatacgcgttgcgtatcgaactcgtcgaagcgttgcgtatgaa 1038
 Db 3550 GTAGAGTTAAATAAATAAGATTAATGTTAGCTACTAATGATCAGGACGAGGCTAGAA 3609
 Qy 1039 aagatctcgtcgtcgtgaagaactcgcgaagcgttgcgtatcgtcgaagcgttgcgtatg 1098
 Db 3610 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3669
 Qy 1099 aatacgaagaacacgaagcgtatcactacaactcgaagaactcgtcgaagaacacat 1158
 Db 3670 AAGTCATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3729
 Qy 1159 ggttaacgatacgttgcgtatcgttgcgaagaactcgaagaactcgtcgaagaactcgt 1218

Db 3730 GGAATGATATAGCTTTATAGATTTCAATTCATTAAATATAGTAAACTAGACA 3789
 Qy 1219 tccaaactgtaacatcgtgaagaactcctcgtcgaactcgttgcgtcgtgag 1278
 Db 3790 ACTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3849
 Qy 1279 tccatccgcgttgcgaagcgttgcgaagcgttgcgaagcgttgcgaagcgttgcga 1320
 Db 3850 TTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3891

RESULT 12
 AX036248
 DEFINITION Sequence 27 from Patent Epi041149.
 ACCESSION AX036248
 VERSION AX036248.1 GI:11225866
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 REFERENCE
 1 (bases 1 to 3891)
 AUTHORS
 Kink,J.A., Firca,J.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and
 Williams,J.A.
 TITLE
 Vaccine and antitoxin for treatment and prevention of C. Difficile
 disease
 JOURNAL
 Patent: EP 1041149-A 27 04-Oct-2000;
 PHARM PHARM INC (US)
 FEATURES
 source
 location/Qualifiers
 1..3891
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..3891
 /note="synthetic"
 /codon_start=1
 /protein_id="CAC16480.1"
 /translation="MOFVKNQFNRKDPVNGVDIAYIKIPNGOMQPVKAFIKHNIW
 IPEBDTPNPREGDLNPPPEAKQVYVSYYDSYISTDENKQNTYKQTKLEPIRYSTD
 IGRMLTSTVKGIPWGGSTIDTELKVIDNCINVTIDPDSYRSEELVITIGSADI
 IOFEKSGHEVLNLTTRNGYSTGYIRSPDPTGFEESELEVEFEELRFGGDAFIDS
 VTNHELTHAGRLVGLIATINPNRVKVTNAYEWSGLEVEFEELRFGGDAFIDS
 LOENEFRLYYNKEFDIASITNKASISYGTASLOQYKMNVEKREKILSEDSGFSVD
 KLRPKILKMLREIYTEDNPVFPFKVRLKRYLNDKAVFKINIVPKVNTIYDGFNL
 RNTNIAANPNQONTETINMNFETKLKNEFGLFEFKLLCVRGLITSKTSLSKQYKAL
 NDCLKVNMDLFFSPEDNFTNDLNKGEELTSDNIEAEENISLDLIQYIYTFN
 DNEPNISIEMLSDIIGOLEMPNIEFPNGKYEIDKYTFHRLAQDEHGSRI
 ALTNSVNEALNPSRVYTFPSSDYVKKNKATKAMPLGAVLEOLVYDPTDESEVST
 DKIDITFIIRYIGPALNIGMLYKDEDEYVGLIFSGAVILLEFPEIRIPIVGFALV
 SYANKVLTVOITDINALSKRNEKDEYVGLIYTNMLAVNOIDILIRKKMEALENOA
 EATKALITVOYNTREEKNININIDLSKINESTIKKAMINIKPLNQSYSILAM
 SMIPGVRLDEPDNSLKDALKYIYDRGTLIGVDRLKRVNNTSTDIYFOLSKY
 VDNORLSTFEYIKNIINTSILNRYSNHLIDLSRAKINIGSKVNPIDIKNOI
 OLFNLESEKIEVILKALIVNSMYENESTSWIRIPKFENSILNNEYTIINCENNS
 GKVSLNNGEILMTODROEIKORVFEYSOMINISDIYINFIPTIYNNLNNSKIY
 INGRILDKPISINGIHASNNIMKEDGCDTRHYIRIKYFNPEDKELNKEIKDY
 DNGNSGILKPFMDYDIODKPYTHYLNIDPRKNYDVANNGIRGTHYIKRGSGVPTT
 NIVLNSSLYRGTKFLIKRYVAGKNDNIYRNDRYIINVVANKERYLATNSOGVEK
 ILSALEIPDVGNLSQVVMKSKNDGINTKCKMIDONNGNDIGFIQHPNNIATKLV
 ASNNYNROIERSRSLGCSWEFIPVDGMBRPL"

BASE COUNT 1580 a 392 c 1290 t
 ORIGIN

Query Match 54.8%; Score 726; DB 6; Length 3891;
 Best Local Similarity 72.4%; Pred. No. 3.1e-176;
 Matches 942; Conservative 0; Mismatches 360; Indels 0; Gaps 0;
 Qy 19 ttactctcgtgaatcgaagaacatcactacactcctcgtgaactcgtgagctac 78
 Db 2590 TCTACATTACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2649

QY 363 gaactacggttgaaatcaltctggaactctgcaggacactcaggaaatcaaacagcgtgtgt 422
DB 386 GAACCTACACAAATCATCTGAGCTCTGAGGACACTGCTGTAACACACAGAACTGCT 445
QY 423 attcaaatctctcagaatgaatcaacatctctgaactacatcaatcgcgtggaatctcggttac 482
DB 446 TTTCAACTACACTCAGATGATCTCTGACTACATTAATAATGATGATCTGCTGATC 505
QY 483 catcaccaaaatctgctgaatcaactccaatctacatcaacgcgcgtctgatcgacca 542
DB 506 TATCACTAACACCGCTGAGTACTCTGTAATCTACATCAACGGTAACCTGATGATGA 565
QY 543 gaaacggatctccaatctcgggtgaacatccacgcctcttaataacatgatgtcaaatgga 602
DB 566 AAAATCTATCTCTAATCAGGCTGAGTACATCCAGCTTCTGACAAATCTCTTCAAAATGCT 625
QY 603 cgggtgtcgtgaacactcaccgctacatctgatacaaatctcaatctgttcgacaaga 662
DB 626 TGGTTGCAACGACAC---GCGTTACGTTGGTATCCGTTACTTCAAAAGTTTTCGACACTGA 682
QY 663 actgaacgaaatgaatcaaacagacctgtagacaacacgcgtccaatctcgtatctgaa 722
DB 683 ACTGGGTAAACCTGAATCGAATCTGACTCTGACGAACCGACCGCTTATCTCTGAA 742
QY 723 agactctcgggtgactacactcgaagtaacgaacacgctactacatgctgaatctgacga 782
DB 743 AGACTTCTGGGGTACTACTGCTGTAACAAACGTTTACTACTGCTGACCTGCTCG 802
QY 783 tccgaacaaatcgttgacgtcaacaatgtaagtaacgcggtttacatgtaccctgaag 842
DB 803 GACTGACAAATCTATCACTCAACACTCT-----AACTTCTGAACTCAACCA 850
QY 843 tccggtgtgtctgtatgactacgaacatctacactcctgacccgtgacggtgtac 902
DB 851 GCAGCGTGTGTTTATCGAATACTTAATATCTTCTTACACTCGTCTGACACTGCTGT 910
QY 903 caaatlcatlcaaaagaatacgcgtct-----ggtacaagaacataatcgttcg 953
DB 911 TGAAGTATCATCCGTAATAAAGGTTTCTACTGACATCTCTAAGACTGACACTGCTGTAG 970
QY 954 caacaatgatcgtgtatatacatcaatgtgtgtaagaaagaataacacgctcgtgtac 1013
DB 971 TAAAAACGACCTGGCTTACATCAACGTTGTGACCGTGAATACCGTCTGTACGC 1030
QY 1014 caatgctctcaagcgtgtgtagaaagaatctgtcgtcgtgaaatccggagcttg 1073
DB 1031 TGACATCTCTATCGCTAAACCGGAAAAATCATCAACTGATCCGTAATCTTA 1090
QY 1074 taa---tctgtctcaggtagtgtgtaatgaatccaagaacgacggtatcactaaca 1130
DB 1091 CAACCTCTCTGGGTGATCATCTGATGACCTC-----GATCGGTACAA 1135
QY 1131 atgcaaatgaatctcgaaggaaacaatgtaacgatatcggtttcaatcgtgttccaca 1190
DB 1136 CTGCACTATGAACCTTCAGAAACACACGCTGTGTAACATCGCTGCTGCTTCACATC 1195
QY 1191 gttcaacaatatcgtaaacatggtcttcacacgtgtaacaatcgtcagatcgaacgttc 1250
DB 1196 TAACAAAC-----CTGTTGCTTCTTCTTGTGTAATCAACATCCGTAATAA 1243
QY 1251 ctctcgacactcgtgtgtcctcttggaagttcatcccggttgatgaaggttggttgaa 1308
DB 1244 CACTTCTTCTTAACGCTGCTCTGCTTCTTCAATCTTAAGAACAACGCTTGGCAGAA 1301

Search completed: September 16, 2002, 21:30:36
Job time: 15628 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 19:28:11 : Search time 4365.92 seconds
(without alignments)
4089.968 Million cell updates/sec

Title: US-09-611-419a-3

Perfect score: 1323

Sequence: 1 gaatcgaacagatgctac.....aacgtccgtctaagaattc 1323

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58.4	4.4	608	10	BI887904 2F637-1-0
2	54.6	4.1	1101	12	AL068607 Drosophila
3	53.6	4.1	537	10	BM277996 AS_tgz-51
4	53.6	4.1	551	10	BM278689 AS_tgz-66
5	53.2	4.0	394	9	AU060224
6	52.6	4.0	447	10	BI594803 AS_tgz-44
7	52	3.9	597	10	BM278181 AS_tgz-54
8	51.8	3.9	500	10	BM278805 AS_tgz-67
9	50.6	3.8	511	10	BM004465 TGESTY2A8
10	50.4	3.8	569	10	BM278657 AS_tgz-65
11	50	3.8	641	10	BM181884 fvs1b1.y
12	49.6	3.7	645	9	AI389106
13	49.2	3.7	827	12	CNS02156
14	49	3.7	481	10	BF050030
15	48.8	3.7	628	9	AU060230
16	48.2	3.6	546	10	C22974
17	48.2	3.6	1101	12	CNS0100X

18	48	3.6	546	10	BF169335
19	48	3.6	556	10	BM277853
20	47.4	3.6	377	10	BI594814
21	47.4	3.6	450	9	AU060996
22	47	3.6	458	10	BI783045
23	47	3.6	518	10	BI782765
24	46.8	3.5	1175	10	BI872945
25	46.4	3.5	469	10	BF050033
26	45.4	3.4	386	10	BF050064
27	45.4	3.4	439	10	BM278692
28	45.4	3.4	536	10	BF050073
29	45.4	3.4	542	10	BM278786
30	45.4	3.4	554	10	C23753
31	45.2	3.4	501	10	BM277916
32	45.2	3.4	525	12	BI184460
33	45.2	3.4	525	12	CNS07PEA
34	45.2	3.4	1101	12	CNS000D1
35	45	3.4	564	10	BE060733
36	44.8	3.4	534	10	BM278087
37	44.6	3.4	1101	12	CNS016JY
38	44.4	3.4	522	10	BM278558
39	44.4	3.4	681	12	CNS02EOD
40	44.2	3.3	325	12	CNS03PK8
41	44.2	3.3	429	9	AU052930
42	44.2	3.3	500	12	B67199
43	44.2	3.3	519	10	BM278164
44	43.8	3.3	535	10	BI594948
45	43.6	3.3	443	10	BM039802

ALIGNMENTS

RESULT 1

BI887904

LOCUS

2F637-1-002159 zebrafish shield stage whole embryo cDNA library

MPMGp637 Danto rerio cDNA clone MPMGP637_10F4:MPMGp637F0410 5', mRNA sequence.

ACCESSION

BI887904.1 GI:16095175

KEYWORDS

SOURCE

ORGANISM

Danio rerio
zebrafish.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 608)
Clark, M., Anstad, P., Hennig, S., Johnson, S.L. and Lehrach, H.

EST sequencing of a zebrafish shield stage cDNA library normalised by oligonucleotide fingerprinting

JOURNAL

COMMENT

Unpublished (2001)
Contact: Hennig S
Laboratory 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Imestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de

5' EST sequencing of clones from a zebrafish shield stage library, normalised from 55,000 starting clones by oligonucleotide fingerprinting

FEATURES

SOURCE

location/Qualifiers
1..608
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="MPMGp637_10F4:MPMGp637F0410"
/clone.lib="zebrafish shield stage whole embryo cDNA library MPMGP637"
/tissue_type="whole embryo"
/dev_stage="shield stage, 6 hrs post-fertilisation"
/lab_host="E.coli, X11 blue MRF"

BASE COUNT	207 a	218 c	41 g	140 t	2 others
ORIGIN	/note="Vector: Spport1; Site 1: NotI; Site 2: SalI; oligo-dr NotI primed. SalI adaptors, directionally cloned library normalised by oligonucleotide fingerprinting"				

Query Match	4.4%	Score 58.4;	DB 10;	Length 608;
Best Local Similarity	49.1%;	Pred. No. 0.00033;		
Matches 181; Conservative	0;	Mismatches 187;	Indels 1;	Gaps 1.

QY	439	atcaacatcctcgcagtaacataca-ccgcctggatcttcgttaccatccacaaatcgtct	497
Db	5	ATCAACTTCTCCAGGTACTACATCACTTCTCCAGCTACATCAACTTCTCCAAACAACA	64
QY	498	gaataactccaanaatcatcatcaacggccgctctgtatcgacccaagaacgattccaatc	557
Db	65	CAACAACATTCTCCAGCTACATCACTTCTCCAAACAACAACAACAACTTCTCCAGCTA	124
QY	558	gggtacaatccacgcgttctaataacatacgaatgttcaaacctggaagcgttgcgtgaacatca	617
Db	125	CATCAACTTCTCCAGCTACATCACTTCTCCAGCTACATCACTTCTCCAGGTACAACA	184
QY	618	ccgcgtacatctggaatcaataacttaataatctgttcgacaagaagcagaagaanaagaat	677
Db	185	CTTCTCCAGCTACATCACTTCTCCAAACAACAACAACAACAACTTCTCCAAACAACA	244
QY	678	caaaagacctgtgaagaacaacagatcccaatctctgtatctccgaagaagctctcggggtgacta	737
Db	245	CAACAACATTCTCCAGCTACATCACTTCTCCAAACAACAACAACAACAACTTCTCCAGCTA	304
QY	738	ccctgcagtagcagaacaacggtactactacagctgaatctctgaagatccggaacaaatacgttga	797
Db	305	CATCAACTTCTCCAGCTACATCACTTCTCCAGCTACATCACTTCTCCAGCTTCTCCAG	364
QY	798	cgtaacaa 806	
Db	365	CTACAACA 373	

RESULT	2
CNS00100	
LOCUS	1101 bp DNA linear
DEFINITION	Drosophila melanogaster genome survey sequence TE73 end of BAC: BACR3D23 of RPci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL068607
VERSION	AL068607.1 GI:4958689
KEYWORDS	GS.
SOURCE	fruit fly
ORGANISM	Drosophila melanogaster Euarysta, Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 1101) Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
JOURNAL	- Web : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaton Mammossor in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2, cn bw sp, the same strain used for the BDGP P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or
COMMENT	

filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

location/Qualifiers

1. .1101

orientation="BACprohibitin_mal3proectoray"

	/note=enu : 1c13			
BASE COUNT	155 a	166 c	7 g	284 t
ORIGIN				
				489 others

Query Match	4.1%	Score 54.6;	DB 12;	Length 1101;
Best Local Similarity	17.6%	Pred. No. 0.0039;		
Matches 100;	Conservative 219;	Mismatches 245;	Indels 3;	Gaps 1.

[illegible]

RESULT	3
LOCUS	BM277996
DEFINITION	BM277996 537 bp mRNA linear EST 20-DEC-2001
ACCESSION	As_tg2_51F02-SKPL Ascaris suum adult male testis germinal zone from
VERSION	Alam.Scot Ascaris suum cDNA clone As_tg2_51F02 5', mRNA sequence.
KEYWORDS	BM277996
SOURCE	BM277996.1 GI:17971254
ORGANISM	EST.
REFERENCE	p1g roundworm.
AUTHORS	Ascaris suum
	Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
	1 (bases 1 to 537)
	Blaxter, M.L., Parkinson, J., Wilton, C., Daub, J., Gulliano, D., Hall
	, N., Quayle, M. and Barrell, B.

TITLE Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk

The library was prepared by Michelle Lizotte-Maniowski for Alan Scott, Johns Hopkins University Medical School, Baltimore. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).

PCR primers
FORWARD: T3
BACKWARD: T7PL
Plate: 51 row: F column: 02
Seq primer: SKPL
High quality sequence stop: 478.
Location/Qualifiers

FEATURES

source

1..537
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_lib="t9z_51F02"
/clone_lib="Ascaris suum adult male testis germinal zone from Alan Scott"
/sex="Male"
/dev_stage="Adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI; Library was made from dissected testis germinal zone from adult male Ascaris suum collected from abattoirs. Constructed by Michelle Lizotte-Maniowski for Alan Scott, Johns Hopkins University Medical School, Baltimore, MD."

BASE COUNT 193 a 149 c 72 g 123 t
ORIGIN

Query Match 4.1%; Score 53.6; DB 10; Length 537;
Best Local Similarity 48.6%; Pred. No. 0.0053;
Matches 180; Conservative 0; Mismatches 184; Indels 6; Gaps 1;

OY 359 tgaactcgtggaatcatctggaactctgcaggaactcgaagaatcaaacagcgtgtg 418
DB 49 tGTACACCAATGCTGCAACAGAACCCACCTGACAACTTCACACAGCAACTTTT 108
OY 419 tatcaaatcactcagatgatcaacatctcgtactaatcaatcgtcgtgactctg 478
DB 109 GCAACTTCACAGTAACCTTCACACAGCAAGCCATTTTCACAGCAACTTCAC 168
OY 479 ccataccacaacatcgtctgaataactccaatactacatacgaagcgctgacgacc 538
DB 169 ACAACCTTAGCAACCTCTGACAACTTCATTTTTCGCAACTCTCTGAGACACG 228
OY 539 agaaacgcatcccaatctgggttaacatccacgctcttaataacatcattcaact 598
DB 229 GGAG-----CTACCAACTTGAACAACCTCAACACCTCAATTAACAACCTGA 282
OY 599 aacgtgtcgagacactacgcgtacatcgtatgaataacttaacttaactcgttcg 658
DB 283 ACAACAATCTCAGCAATCTCAACAATTAACCTTCACAACTTGAACAGCAACTTCA 342
OY 659 aactgaagaaagaatacaagaccgtgacgaacacagtcgaattcgtgtatcctga 718
DB 343 ACAGTAATCTCAGCAACTTCACACAGCAACAACTTGAAGCAACTTCAACAAGTA 402
OY 719 aagactctg 728
DB 403 ACAACAGCAG 412

RESULT 4
BM278689

LOCUS BM278689 551 bp mRNA linear EST 20-DEC-2001
DEFINITION As.t9z.66A05.SKPL Ascaris suum adult male testis germinal zone from Alan Scott
ACCESSION BM278689
VERSION BM278689.1 GI:17971947
KEYWORDS EST.
SOURCE pig roundworm.
ORGANISM Ascaris suum
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea

REFERENCE 1 (bases 1 to 551)
Blaxter,M.L., Parkinson,J., Wilton,C., Daub,J., Guiliano,D., Hall,N., Quayle,M. and Barrell,B.
Edinburgh University/Sanger Centre Nematode EST Project
Unpublished (2000)

AUTHORS

TITLE

JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk

The library was prepared by Michelle Lizotte-Maniowski for Alan Scott, Johns Hopkins University Medical School, Baltimore. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).

PCR primers
FORWARD: T3
BACKWARD: T7PL
Plate: 66 row: A column: 05
Seq primer: SKPL
High quality sequence stop: 510.
Location/Qualifiers

FEATURES

source

1..551
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_lib="t9z_66A05"
/clone_lib="Ascaris suum adult male testis germinal zone from Alan Scott"
/sex="Male"
/dev_stage="Adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI; Library was made from dissected testis germinal zone from adult male Ascaris suum collected from abattoirs. Constructed by Michelle Lizotte-Maniowski for Alan Scott, Johns Hopkins University Medical School, Baltimore, MD."

BASE COUNT 208 a 160 c 64 g 119 t
ORIGIN

Query Match 4.1%; Score 53.6; DB 10; Length 551;
Best Local Similarity 49.6%; Pred. No. 0.0053;
Matches 137; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

OY 428 acttccagatgataacatctcgtactacataatcgcgtgactctcgttaccatcaca 487
DB 102 ACAGTAATCTCAACAACGATTTTCACAGCAAGCACTGACAACTTCATACACCT 161
OY 488 acaatcgtctgataatactcaaatatcatcaatcaacgacgctctgacgaaacgga 547
DB 162 GCAACTCTCTGACAACTTCATTAACAACCTGACAACTCTGAGCAACCTCACTCA 221
OY 548 tctcaatcctgggttaacatccacgctctcaataacatcatcgttcaactcgtgctg 607
DB 222 ACTTGACAACTTCACAACTTCACAACTTCACAACTTCACAACTTCACAACTTC 281
OY 608 gtgacactcaacgctacatctgatacaataatcattcaatcgttcgacaagaactgac 667
DB 282 GTCTCAACAATTAACCTTCACAACTTCAGCAACGACGAAATTTTCACAGCAAGTA 341
OY 668 aaaaagaatcaagacctgtgacgaacacgactga 703

ACCESSION Alan Scott Ascaris suum cDNA clone As_tgz_54C05 5', mRNA sequence.
 BM278181
 VERSION BM278181.1 GI:17971439
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.
 1 (bases 1 to 597)
 Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D., Hall
 N., Quayle,M. and Barrell,B.
 Edinburg University/Sanger Centre Nematode EST Project
 Unpublished (2000)
 TITLE Contact: Blaxter ML
 JOURNAL Institute of Cell, Animal and Population Biology
 COMMENT University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3JF, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared by Michelle Lizotte-Waniewski for Alan
 Scott, Johns Hopkins University Medical School, Baltimore.
 Sequencing was performed by the Pathogen Sequencing Unit, Sanger
 Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).
 PCR Primers
 FORWARD: T3
 BACKWARD: T7PL
 Plate: 54 row: C column: 05
 Seq primer: SKPL
 High quality sequence stop: 502.
 Location/Qualifiers

FEATURES
 source
 1..597
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone="As_tgz_54C05"
 /clone_lib="Ascaris suum adult male testis germinal zone
 from Alan Scott"
 /sex="Male"
 /dev_stage="Adult"
 /note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
 library was made from dissected testis germinal zone from
 adult male Ascaris suum collected from abattoirs.
 Constructed by Michelle Lizotte-Waniewski for Alan Scott,
 Johns Hopkins University Medical School, Baltimore, MD."

BASE COUNT 222 a 167 c 76 g 132 t
 ORIGIN

Query Match 3.9%; Score 52; DB 10; Length 597;
 Best Local Similarity 49.3%; Pred. No. 0.014;
 Matches 136; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

428 actctcagatgatcaacatctctgactacatcaatcgctgtagcttcgttaccatcacca 487
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 141 ACAGTAACTTCAACACCAATTTCACAGCAACTTCACATTAACACCTTA 200
 488 acaatcgctgtagtaactcaaatctacatacaagccgctgtagtcgacgaacgca 547
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 201 GCAACTCTCTGAACAACCTTCAATACCACTCAGCAACTGCTGAGCAACCTCAGTTTCA 260
 548 tctccaatctgggttaactcaacgcttcaatacatcatggttcaactgagcggtgtgc 607
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 261 ACTTGAACAACCTCAACACCTTCATACCACTCAACTATCTCAACAACAATTCAGCA 320
 608 gtgacactcaacgctacatctgtagtaacatactcaatctgttcgacaagaactgaag 667
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 321 ATCTCAACAATTAACCTCAACACCTTACGACGAGCAACTTCAACGACGTAATCTCAACA 380
 668 aaaaagaataaagacactgttagcaaacacgtcca 703
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 381 ACAGTAACTTCAACCACTTCAACGACCAACTTCA 416

RESULT 8
 LOCUS BM278805 500 bp mRNA linear EST 20-DEC-2001
 DEFINITION As_tgz_67F10.SKPL Ascaris suum adult male testis germinal zone from
 Alan Scott Ascaris suum cDNA clone As_tgz_67F10 5', mRNA sequence.
 ACCESSION BM278805
 VERSION BM278805.1 GI:17972063
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.
 1 (bases 1 to 500)
 Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D., Hall
 N., Quayle,M. and Barrell,B.
 Edinburg University/Sanger Centre Nematode EST Project
 Unpublished (2000)
 TITLE Contact: Blaxter ML
 JOURNAL Institute of Cell, Animal and Population Biology
 COMMENT University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3JF, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared by Michelle Lizotte-Waniewski for Alan
 Scott, Johns Hopkins University Medical School, Baltimore.
 Sequencing was performed by the Pathogen Sequencing Unit, Sanger
 Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The
 sequence contained a Polya tail (trimmed)
 PCR Primers
 FORWARD: T3
 BACKWARD: T7PL
 Plate: 67 row: F column: 10
 Seq primer: SKPL
 High quality sequence stop: 500.
 Location/Qualifiers

FEATURES
 source
 1..500
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone="As_tgz_67F10"
 /clone_lib="Ascaris suum adult male testis germinal zone
 from Alan Scott"
 /sex="Male"
 /dev_stage="Adult"
 /note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
 library was made from dissected testis germinal zone from
 adult male Ascaris suum collected from abattoirs.
 Constructed by Michelle Lizotte-Waniewski for Alan Scott,
 Johns Hopkins University Medical School, Baltimore, MD."

BASE COUNT 193 a 144 c 58 g 105 t
 ORIGIN

Query Match 3.9%; Score 51.8; DB 10; Length 500;
 Best Local Similarity 50.8%; Pred. No. 0.015;
 Matches 153; Conservative 0; Mismatches 142; Indels 6; Gaps 1;

428 actctcagatgatcaacatctctgactacatcaatcgctgtagcttcgttaccatcacca 487
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 70 ACAGTAACTTCAACACCAATTTCACAGCAACTTCACATTAACACCTTA 129
 488 acaatcgctgtagtaactcaaatctacatacaagccgctgtagtcgacgaacgca 547
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 130 GCAACTCTCTGAACAACCTTCAATACCACTCAGCAACTGCTGAGCAACCTCAG----- 184
 548 tctccaatctgggttaactcaacgcttcaatacatcatggttcaactgagcggtgtgc 607
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 Db 185 -CTACAACTTGAACAACCTTCAACCACTTCAACCACTTCAACTATCTCAACAACAATTC 243
 608 gtgacactcaacgctacatctgtagtaacatactcaatctgttcgacaagaactgaag 667
 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11


```

/sex="Male"
/dev_stage="Adult"
/notes="Vector: Lambda Zap II; Site 1: EcoRI; Site 2: XhoI;
Library was made from dissected testis germinal zone from
adult male Ascaris suum collected from abattoirs.
Constructed by Michelle Lizotte-Waniewski for Alan Scott,
Johns Hopkins University Medical School, Baltimore, MD."
BASE COUNT      213 a      164 c      69 g      123 t
ORIGIN

```

```

Query Match      3.8%; Score 50.4; DB 10; Length 569;
Best Local Similarity 48.9%; Pred. No. 0.035;
Matches 135; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY  428 attctcaatgatcaactctctctgactatcatcaatcgctgactctgttaccatcacca 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  119 ACAGTAACTTCAACACAGCAATTTCAACAGCAACCTTCAAGCAATTAACAACCTTCA 178
QY  488 acaatcgctgataactccaaatctacatcaacgagcgctctgactgacgagaaccca 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  179 GCAATCTCTGACACACTTCAATACACACTGACAACTCTGTGACAACTCAGACTACA 238
QY  548 tctccaatctggtacatccacgcttcttaatacatcatgtaactgagcaggtgttc 607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  239 ACTTGAAACACCTCAACACACCTCAATACACCTCAATCTCAACACACATCTCAGCA 298
QY  608 gtgacactcaccgctacatctggtacaaatctcaatctgttgcagaagaactgaacg 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  299 ATCTCAACAAATACCTTCAACAACTTACAGCAGCAACTTCAACAGCAAGTAACTTCA 358
QY  668 aaaaagaatcaagaactgtgacgacacccagttcca 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  359 ACAGTAACTTCAACACACTTCAACACAACTTCA 394

```

```

RESULT 11      641 bp  mRNA  linear  EST 11-DEC-2001
BM181884
LOCUS      f951b11.y1 Sugano STD adult male Dario rerio cDNA clone 5412044.5
DEFINITION  similar to contains element TARI repetitive element ;, mRNA

```

```

ACCESSION  BM181884
VERSION    BM181884
KEYWORDS   zebrafish.
SOURCE     Dario rerio
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Dario.
            1 (bases 1 to 641)

```

```

REFERENCE
AUTHORS   Clark M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
            , S., Hillier, L., Kucaba, T., Martin, U., Beck, C., Wylie, T., Underwood
            , K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
            Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E.,
            Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.,
            and Wilson, R.

```

```

TITLE      Mashu zebrafish EST project 1998
JOURNAL    Unpublished (1998)
COMMENT    Contact: Stephen L. Johnson
            Washington University School of Medicine
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810

```

```

FEATURES
SOURCE     Email: zbrfish@wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution information can be found through the I.M.A.G.E.
            Consortium/UNL, send email to: info@image.llnl.gov
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 527.
            Location/Qualifiers
            1..641

```

```

/organism="Dario rerio"
/db_xref="taxon:7955"
/clone_id="5412044"
/clone_lib="Sugano STD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/notes="Vector: pME185-FL3; Site 1: DraIII (CACCATGNG);
Site 2: DraIII (CACCATGNG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGggccctttttttttttttt];
double-stranded cDNA was ligated to a DraIII adaptor
(TGTTGGCCCTACTGG), digested and cloned into distinct DraIII
sites of the pME185-FL3 vector (5' site CACGTGNG, 3' site
CACCATGNG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTGTGCTTAAAGTCGCG and 3' end primer
CGACCTCGACCTCGAGCA."
BASE COUNT      218 a      234 c      37 g      152 t
ORIGIN

```

```

Query Match      3.8%; Score 50; DB 10; Length 641;
Best Local Similarity 46.3%; Pred. No. 0.047;
Matches 201; Conservative 0; Mismatches 230; Indels 3; Gaps 1;

```

```

QY  226 tacaactcatgtacaaactctccactctctctgactcgtatccggaataacttc 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  175 TACATCAACTTCTCCACACACAAACAACTTCTCCACACAAACAACTTCTCCAGTACATC 234
QY  286 aactcatctctctgacaaatgaataaccatcatcaactgacatggaataaa---ttct 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  235 AACTTCAACACAAACAAACAACTTCTCCAGTACATCACTTCAACAAACAAACAACTTCTCC 294
QY  343 gttgagaagtatctctgactgagtggaatcatctgactctgacaggaactgaaga 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  295 ACTTCAACAACTTCTCCAGCACAACATCACTTCTCCAGTACATCACTTCTCCACAAC 354
QY  403 atcaaacagcgtgtgtatcaataactctcagatgatcaatctctgactacatcaat 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  355 AACCACTTCTCCAGTACATCACTTCTCCACAACAAACAACTTCTCCAGTACATCAAC 414
QY  463 cgcgtgactctctgtacatccacaaatcgctcgtgaataactccaatactacatcaac 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  415 TTCTCCAAACAAACAACTTCTCCAGTACATCACTTCAACAAACAAACAACTTCTCC 474
QY  523 ggcgctgactgacaggaacacgactctcaatctcgtgtaaatccagcttctaataac 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  475 ACCTACATCACTTCAACACAAACAACTTCTCCAGTACATCACTTCTCCAGCACAACATC 534
QY  583 atcatgttcaaaactgacggtgtcgtgacacacacgctacatctgataactacatc 642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  535 AACTTCTCCAGTACATCACTTCAACAAACAAACAACTTCTCCAGTACATCAACATTC 594
QY  643 aatctgctgcagaa 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  595 AACCAACAAACAAACAA 608

```

```

RESULT 12      645 bp  mRNA  linear  EST 23-APR-2001
AI389106
LOCUS      GH20192.5prine GH Drosophila melanogaster head p0r2 Drosophila
DEFINITION  melanogaster cDNA clone GH20192.5, mRNA sequence.
ACCESSION  AI389106
VERSION    AI389106.2 GI:13758704
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```


BF050030 481 bp mRNA linear EST 16-OCT-2000
LOCUS As.tgz_21C12.SKPL Ascaris suum adult male testis germinal zone from
DEFINITION Alan Scott Ascaris suum cDNA clone As.tgz_21C12.5', mRNA sequence.
ACCESSION BF050030
VERSION BF050030.1 GI:10803926
KEYWORDS EST.
SOURCE pig roundworm.
ORGANISM Ascaris suum
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
; Ascarididae; Ascaris.
REFERENCE 1 (bases 1 to 481)
AUTHORS Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guiliano, D., Hall
N., Quayle, M. and Barrell, B.
TITLE Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL Unpublished (2000)
COMMENT Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 3450
Email: mark.blaxter@ed.ac.uk
The library was prepared by Michelle Lizotte-Waniewski for Alan
Scott, Johns Hopkins University Medical School, Baltimore.
Sequencing was performed by Claire Whitton ICRP, Edinburgh
PCR primers
FORWARD: T3
BACKWARD: T7PL
Plate: 21 row: C column: 12
Seq primer: SKPL
High quality sequence stop: 456.
FEATURES
source location/Qualifiers
1..481
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone="As.tgz_21C12"
/clone_lib="Ascaris suum adult male testis germinal zone
from Alan Scott"
/sex="Male"
/dev_stage="Adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
Library was made from dissected testis germinal zone from
adult male Ascaris suum collected from abattoirs.
Constructed by Michelle Lizotte-Waniewski for Alan Scott,
Johns Hopkins University Medical School, Baltimore, MD."
BASE COUNT 188 a 149 c 49 g 95 t
ORIGIN
Query Match 3.7%; Score 49; DB 10; Length 481;
Best Local Similarity 49.1%; Pred. No. 0.075; Indels 0; Gaps 0;
Matches 130; Conservative 0; Mismatches 135;
439 atcaacatctcgactacaatcgatcgatctcgatcattacatcaccaaatcgctg 498
104 AACAAAGCAATTTTCACGACCACTTCACCACTTCAATTAACAACTTGCAACTCTTG 163
499 aataatccaaaatcacatcaacggcgctgctgacgcagaaacgaatctccaatctg 558
164 AACAACTTCAATTAACAACTTCAGCAACCTCTGAGCAACCTGACCTAAGTAAACAAC 223
559 ggtaacatccagctcttcatataacatcatgatcgaactgagcgtgtgtgacatcgc 618
224 CTCAACAACCTCAATTAACAACTTCAATTCACCAATCTCAGCAATCTCAACAAT 283
619 cgtacatctgagatcaaatctcaatctgttcgcagaaagaaactgaaagaaatc 678
284 AACCTCAACAACCTTACGACGCAACTTCAACAGCACTACTTCAACAACAGTACTTTC 343
679 aaagactgtacgacaacacatcca 703

Db 344 AGCAACTTCACAGCAACAACTTCA 368
RESULT 15
LOCUS AU060230 628 bp mRNA linear EST 20-MAY-1999
DEFINITION AU060230 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
discoideum cDNA clone SLA628, mRNA sequence.
ACCESSION AU060230
VERSION AU060230.1 GI:488134
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
REFERENCE 1 (bases 1 to 628)
AUTHORS Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,
Yoshino, R., Mita, B.N., Pl.M., Sato, T., Takemoto, K., Yasukawa, H.,
Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
TITLE Developmental cDNA in Dictyostelium discoideum
JOURNAL Unpublished (1998)
COMMENT Contact: Hideto Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = Dictyostelium discoideum cDNA project in Japan.
FEATURES
source location/Qualifiers
1..628
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLA628"
/clone_lib="Dictyostelium discoideum SL (H.Urushihara)"
/dev_stage="slug"
BASE COUNT 270 a 160 c 56 g 142 t
ORIGIN
Query Match 3.7%; Score 48.8; DB 9; Length 628;
Best Local Similarity 48.6%; Pred. No. 0.094; Indels 0; Gaps 0;
Matches 134; Conservative 0; Mismatches 142;
423 caaatctcagatgatcaacaatctctgactacatcaatcgatcgatctgcttaccat 482
78 CAACCAACCTTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 137
483 caccacaatcgctgtaataactcaaatctacatcaacggcgctgctgacccagaa 542
138 CAATAACAACCAATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 197
543 accgatccaatctggtgaacatccagctcttaataacatcatgatgttcaaatgagcg 602
198 CAACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 257
603 ttgtcgtagacatccagctgacatctggaatcaatctcaatctgctgacaaagaact 662
258 ATTCAACAACAATATTCACAACAACATCAACAACAACAACAACAACAACAACAACA 317
663 gaacgaagaagaatcaagaacctgtgacgaacca 698
318 CAACAACAACAATATTCACAACAATGTTTATTCACAACAACA 353

Search completed: September 16, 2002, 19:28:15
Job time: 8372 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:43:19 : Search time 631.73 Seconds
(without alignments)
3603.800 Million cell updates/sec

Title: US-09-611-419a-5

Perfect score: 1326

Sequence: 1 gaattcgaaacatgagcttc.....aacgtccgtcgtgaagaattc 1326

Scoring table: IDENTITY_NUC

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	100.0	22	AAA54484	Botulinum toxin hea
2	1310	98.8	22	AAA54483	Botulinum toxin hea
3	1308.2	98.7	22	AAA54482	Botulinum toxin hea
4	1302.2	98.2	17	AAAT9245	Type A neurotoxin
5	1302.2	98.2	19	AAV30571	Clostridium botuli
6	1302.2	98.2	21	AAAT87212	DNA encoding synth
7	1302.2	98.2	1351	AAV30576	Clostridium botuli
8	1302.2	98.2	17	AAAT9246	Type A neurotoxin
9	1302.2	98.2	19	AAV30572	Clostridium botuli

10	1247.8	94.1	1332	22	AAA54588	Sequence encoding
11	727.2	54.8	1546	19	AAV30575	Clostridium botuli
12	727.2	54.8	1435	21	AAAC64582	BONT/A neurotoxin
13	726	54.8	1317	21	AAAT87220	DNA encoding nativ
14	726	54.8	2532	21	AAAT87218	DNA encoding nativ
15	726	54.8	3891	17	AAAT9244	C. botulinum type
16	685.2	51.7	702	21	AAAT87221	DNA encoding BONTA
17	619	46.7	621	21	AAAT87222	DNA encoding BONTA
18	422	31.8	1313	18	AAAT48101	Immunogenic type F
19	422	31.8	1314	22	AAA54499	Botulinum toxin hea
20	400.6	30.2	1317	21	AAAT87216	DNA encoding synth
21	398.4	30.0	1317	22	AAA54490	Botulinum toxin hea
22	371.4	28.0	1917	21	AAAT87219	DNA encoding nativ
23	289.4	21.8	1347	21	AAAT87217	DNA encoding synth
24	289.4	21.8	1368	22	AAA54491	Botulinum toxin hea
25	264.8	20.0	1347	22	AAA54589	Sequence encoding
26	263	19.8	1341	21	AAAT87213	DNA encoding synth
27	263	19.8	1341	22	AAA54485	Botulinum toxin hea
28	261.8	19.7	1278	22	AAA54489	Botulinum toxin hea
29	261.8	19.7	1400	21	AAAT87215	DNA encoding synth
30	261.8	19.7	1400	22	AAA54488	Botulinum toxin hea
31	227	17.1	1293	18	AAAT48100	Immunogenic type F
32	202.8	15.3	1374	22	AAA54487	Botulinum toxin hea
33	193.2	14.6	1472	19	AAV30585	Clostridium botuli
34	193.2	14.6	4017	22	AAAT87212	C botulinum BONT/E
35	190	14.3	1463	19	AAV30584	Clostridium botuli
36	180.8	13.6	1460	19	AAV30593	Clostridium botuli
37	168.8	12.7	1371	21	AAAT87214	DNA encoding synth
38	168.8	12.7	1371	22	AAA54486	Botulinum toxin hea
39	166.4	12.5	1359	12	AAO12121	Synthetic tetanus
40	166.4	12.5	3712	16	AAO97490	Plasmid pHTPA1. N
41	166.4	12.5	3754	15	AAO57880	Intermediate plasm
42	166.4	12.5	3754	16	AAO85420	Plasmid pTECH1. S
43	166.4	12.5	3769	15	AAO57881	Intermediate plasm
44	166.4	12.5	3769	16	AAO97492	Plasmid pTECH2. N
45	166.4	12.5	4366	16	AAO85424	Plasmid pTECH3-P28

ALIGNMENTS

RESULT 1
ID AAA54484
AAA54484 standard; DNA; 1326 bp.
XX
AC AAA54484;
XX
DT 11-APR-2001 (first entry)
XX
DE Botulinum toxin heavy chain C-terminal coding sequence (serotype A).
XX
KW Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
KW recombinant vector; antigen; Immune response; vaccine; Bacterium;
KW Infection; ds.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
FH Key Location/Qualifiers
FT CDS 13..1320 /*tag= a
FT /product= H_C peptide fragment
XX
PN WO200067700-A2.
XX
PD 16-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US12890.
XX
PR 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.

PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.
 XX
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX
 PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 XX
 DR WPI: 2001-016048/02.
 DR P-PSDB: AAB04090.
 XX
 XX New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism
 XX
 PS Disclosure: Fig 3a; 73pp; English.
 XX
 CC Botulinum neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dichain
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.
 XX
 SO Sequence 1326 BP; 404 A; 336 C; 243 G; 343 T; 0 other;

Query Match 100.0%; Score 1326; DB 22; Length 1326;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 1326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaatcgaacagatgacctctactctatcgaatatacatcaagaacatcatactacc 60
 DB 1 gaatcgaacagatgacctctactctatcgaatatacatcaagaacatcatactacc 60
 QY 61 atctggaacctgctgaacgaatccaatcacctgacacgtgtctcgtactgctccaa 120
 DB 61 atctggaacctgctgaacgaatccaatcacctgacacgtgtctcgtactgctccaa 120
 QY 121 atcaacatcggttctaagttaacttgcgtatccgataaagaatcagaatcagctgttc 180
 DB 121 atcaacatcggttctaagttaacttgcgtatccgataaagaatcagaatcagctgttc 180
 QY 181 aatcgaatcttccaanaatcgaaatctacccaagaatgcatcgtatacaactctatg 240
 DB 181 aatcgaatcttccaanaatcgaaatctacccaagaatgcatcgtatacaactctatg 240
 QY 241 aatcgaatcttccaanaatcgaaatctacccaagaatgcatcgtatacaactctatg 240
 DB 241 aatcgaatcttccaanaatcgaaatctacccaagaatgcatcgtatacaactctatg 240
 QY 241 taagaagaacttccactctctctgtagtcglatcccgaaataactccaactctatc 300
 DB 241 taagaagaacttccactctctctgtagtcglatcccgaaataactccaactctatc 300
 QY 301 ctgaacaatgaatacacaatcactgaatgagaanaaatcttggttggaagtatct 360
 DB 301 ctgaacaatgaatacacaatcactgaatgagaanaaatcttggttggaagtatct 360
 QY 361 ctgaactacggtgaataatcatctggaactctgacgagacaacccaagaatcagaatc 420
 DB 361 ctgaactacggtgaataatcatctggaactctgacgagacaacccaagaatcagaatc 420
 QY 421 gtatcaataactctcagatgatacaacatctctgaactacatacaatcgtgtactgt 480
 DB 421 gtatcaataactctcagatgatacaacatctctgaactacatacaatcgtgtactgt 480

QY 481 accatcacaacaatctgctgaataaaccaccaaatctacatacaacgagcgctgtatcgac 540
 DB 481 accatcacaacaatctgctgaataaaccaccaaatctacatacaacgagcgctgtatcgac 540
 QY 541 cagaacccgactcccaatctggttaacatccacgctctcaataaactcaatgttcaacg 600
 DB 541 cagaacccgactcccaatctggttaacatccacgctctcaataaactcaatgttcaacg 600
 QY 601 gacggttcgtagacacatcacgctacatctgataatcaataactcaatctgttcgacaa 660
 DB 601 gacggttcgtagacacatcacgctacatctgataatcaataactcaatctgttcgacaa 660
 QY 661 gaactgaacgaanaaagaatccaagaacctgtacgacacacagtcgaatctgtatccg 720
 DB 661 gaactgaacgaanaaagaatccaagaacctgtacgacacacagtcgaatctgtatccg 720
 QY 721 aaagactctgggtgtaacactgcagctgacgaacacgctacataatgtagtactgtac 780
 DB 721 aaagactctgggtgtaacactgcagctgacgaacacgctacataatgtagtactgtac 780
 QY 781 gatcgaacaataacgctgacgctcaacaatgtagtactgcgagctgtacatgtacgaa 840
 DB 781 gatcgaacaataacgctgacgctcaacaatgtagtactgcgagctgtacatgtacgaa 840
 QY 841 ggtcgcggtgtctgctttagtactccaacatctactctgaactctcctctgacggtgt 900
 DB 841 ggtcgcggtgtctgctttagtactccaacatctactctgaactctcctctgacggtgt 900
 QY 901 accaaatcatcatcaagaataacgctgtgtgaacaagaacataatctgttcgacaact 960
 DB 901 accaaatcatcatcaagaataacgctgtgtgaacaagaacataatctgttcgacaact 960
 QY 961 gatcgtgatacatcaatgctgttagtgaagaacaagaataccgtctgtaccgaatgtct 1020
 DB 961 gatcgtgatacatcaatgctgttagtgaagaacaagaataccgtctgtaccgaatgtct 1020
 QY 1021 tctcagcgtggtgtgaagaatctgtctgtctgtgaatcccgagctgtgtaatctg 1080
 DB 1021 tctcagcgtggtgtgaagaatctgtctgtctgtgaatcccgagctgtgtaatctg 1080
 QY 1081 tctcagatgctgtgaatgaatccaagaacgacgaagatcatcactacaatgcaaatg 1140
 DB 1081 tctcagatgctgtgaatgaatccaagaacgacgaagatcatcactacaatgcaaatg 1140
 QY 1141 aatctgcaagacaacaatgtaacgatalcgtgttcatctgtgttccacagttcaacaat 1200
 DB 1141 aatctgcaagacaacaatgtaacgatalcgtgttcatctgtgttccacagttcaacaat 1200
 QY 1201 atcgcataaactgtgtctccaactggtatacatcgtacgaatggaatggtcctctgcact 1260
 DB 1201 atcgcataaactgtgtctccaactggtatacatcgtacgaatggaatggtcctctgcact 1260
 QY 1261 ctgggtgtctcttggtgagttcatcccggtgtgtagacggttgggtggaagctcgcgttaa 1320
 DB 1261 ctgggtgtctcttggtgagttcatcccggtgtgtagacggttgggtggaagctcgcgttaa 1320
 QY 1321 gaattc 1326
 DB 1321 gaattc 1326

RESULT 2
 AAA54483
 ID AAA54483 standard; DNA; 1323 BP.
 XX
 AC AAA54483;
 XX
 DT 11-APR-2001 (first entry)
 XX
 DE Botulinum toxin heavy chain C-terminal coding sequence (serotype A).
 XX Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;

KM Infection: ds.
 XX Synthetic.
 OS Clostridium botulinum.
 XX
 FH Key Location/Qualifiers
 FT 13..1317
 FT CDS /*tag- a
 /product= H_C peptide fragment
 XX
 PN M0200067700-A2.
 XX
 PD 16-NOV-2000.
 XX
 PF 12-MAY-2000; 2000MO-US12890.
 XX
 PR 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.
 XX
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX
 PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 XX
 DR WPI: 2001-016048/02.
 DR P-PSDB; AAB04089.
 XX
 PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism
 XX
 PS Disclosure; Fig 2a; 73pp; English.
 XX
 XX Botulinism neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dichain
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.
 XX
 XX Sequence 1323 BP; 404 A; 334 C; 242 G; 343 T; 0 other;
 S0

Query Match 98.8%; Score 1310; DB 22; Length 1323;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 gaattcgaaacgattgctctacattacatgaatatacaagaacatcatcaatctcc 60
 Db 1 gaattcgaaacgattgctctacattacatgaatatacaagaacatcatcaatctcc 57
 QY 61 atcctgaacctgctgaatcaatcaatcgaatcgaatcgaatcgaatcgaatcgaat 120
 Db 58 atcctgaacctgctgaatcaatcaatcgaatcgaatcgaatcgaatcgaatcgaat 117
 QY 121 atcaaatcggtcttaagttaactcgaatcgaatcgaatcgaatcgaatcgaatcgaat 180
 Db 118 atcaaatcggtcttaagttaactcgaatcgaatcgaatcgaatcgaatcgaatcgaat 177

QY 181 aatctgaatctcccaaatcgaagtatactctgaagaatgtatcgtatatacaactctatg 240
 Db 178 aatctgaatctcccaaatcgaagtatactctgaagaatgtatcgtatatacaactctatg 237
 QY 241 taagaaactctcccaactcctctgtatccgtatcccgaaatactcaactcaatctct 300
 Db 238 taagaaactctcccaactcctctgtatccgtatcccgaaatactcaactcaatctct 297
 QY 301 ctgaacaatgaatacaccatcatcaatcaactcgaatgaagaatctcgtgtgaagaatgatc 360
 Db 298 ctgaacaatgaatacaccatcatcaatcaactcgaatgaagaatctcgtgtgaagaatgatc 357
 QY 361 ctgaactcgttgaataatcatctgtactcgtcgaagacactcgaagaatacaacagtggt 420
 Db 358 ctgaactcgttgaataatcatctgtactcgtcgaagacactcgaagaatacaacagtggt 417
 QY 421 gtattcaaatctctcagaatgaatcaatctctgtactatcaatcgtgtgattctggt 480
 Db 418 gtattcaaatctctcagaatgaatcaatctctgtactatcaatcgtgtgattctggt 477
 QY 481 accatcaccacaacatcgtctgaataatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 540
 Db 478 accatcaccacaacatcgtctgaataatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 537
 QY 541 cagaacccgaatctcccaatcgtgtatatacaccagctcttaataatcatatgtttcaactg 600
 Db 538 cagaacccgaatctcccaatcgtgtatatacaccagctcttaataatcatatgtttcaactg 597
 QY 601 gacggtgtcgtgaacactcaccgctacatctgtatcaaatcaatcattcattgttcgacaa 660
 Db 598 gacggtgtcgtgaacactcaccgctacatctgtatcaaatcaatcattcattgttcgacaa 657
 QY 661 gaactcgaagaaaagaataataagaactcgtatagacgaacacggtccaatctcgttatctgt 720
 Db 658 gaactcgaagaaaagaataataagaactcgtatagacgaacacggtccaatctcgttatctgt 717
 QY 721 aaagactctcgtgtgactactcgtcagtagacgaacacggtactatcgtatcgtatcgtac 780
 Db 718 aaagactctcgtgtgactactcgtcagtagacgaacacggtactatcgtatcgtatcgtac 777
 QY 781 gatccgaacaatacgttgaatcgaatcaaatgaatgaatgaatgaatgaatgaatgaatgaat 840
 Db 778 gatccgaacaatacgttgaatcgaatcaaatgaatgaatgaatgaatgaatgaatgaatgaat 837
 QY 841 ggtccggtgtgtcgttatgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 900
 Db 838 ggtccggtgtgtcgttatgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 897
 QY 901 accaatcatcatcaatgaagaatacgcgtcgtgtgaacgaaggaataatcgttcgacaat 960
 Db 898 accaatcatcatcaatgaagaatacgcgtcgtgtgaacgaaggaataatcgttcgacaat 957
 QY 961 gatcgttatcatcaatcgttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1020
 Db 958 gatcgttatcatcaatcgttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1017
 QY 1021 tctcaggtgtgtatgaagaataatcgtcgtcgtcgtgaatcccggaatcgtgtgaatctg 1080
 Db 1018 tctcaggtgtgtatgaagaataatcgtcgtcgtcgtgaatcccggaatcgtgtgaatctg 1077
 QY 1081 tctcaggtgtgtatgaagaataatcgaagacgacggtgtatcatcaataatgaagaatg 1140
 Db 1078 tctcaggtgtgtatgaagaataatcgaagacgacggtgtatcatcaataatgaagaatg 1137
 QY 1141 aatctcagaacaacaatcgttaagatatcgtttcattcgtttccaccagttcaaat 1200
 Db 1138 aatctcagaacaacaatcgttaagatatcgtttcattcgtttccaccagttcaaat 1197
 QY 1201 atcgttaaacgtgtgtcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1260
 Db 1198 atcgttaaacgtgtgtcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1257
 QY 1261 ctgggtgtcgtcttggagttcatcccggttgaatgaatcgttgggttgaacgtcgtgtga 1320

|||||
Db 1358 ctgggttgctcttgaggatccatcccggttgagcggctgggggtgaacgtccgtgtaa 1317
Qy 1321 gaattc 1326
|||||
Db 1318 gaattc 1323

RESULT: 3
AAA54482
ID AAA54482 standard; DNA; 1332 BP.
XX

AC AAA54482;
DT 11-APR-2001 (first entry)
XX
DE Botulinism toxin heavy chain C-terminal coding sequence (serotype A).
XX
KM Botulinism toxin; neurotoxin; heavy chain; recombinant expression;
KW recombinant vector; antigen; immune response; vaccine; Dactylum;
KW infection; ds.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
FH Key Location/Qualifiers
FT CDS 13..1326
FT /*tag= a
FT /product= H_C peptide fragment
XX
PN WO200067700-A2.
XX
PD 16-NOV-2000.
XX
PE 12-MAY-2000; 2000WO-US12890.
XX
PR 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.
PR 12-MAY-1999; 99US-0133869.
PR 12-MAY-1999; 99US-0133873.
PR 29-JUL-1999; 99US-0146192.
XX
XX (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
PA
PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX
DR WPI: 2001-016048/02.
DR P-PSDB: AAB04088.
XX
XX New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulinism
XX
XX Claim 2; Fig 1a; 73pp; English.
PS
XX Botulinism neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.

XX SQ Sequence 1332 BP; 404 A; 337 C; 245 G; 346 T; 0 other;
Query Match 98.7%; Score 1308.2; DB 22; Length 1332;
Best local similarity 99.8%; Pred. No. 0;
Matches 1310; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 14 tggcctctactcctcactgaatacatcaagaacatcatcaatacctccatcgtgaacctgc 73
Db 20 tgcgtctcaactcctcactgaatacatcaagaacatcatcaatacctccatcgtgaacctgc 79
Qy 74 gctagaatccaatactaccctgtgcgcctgctcgtctgaacgtctccaataccaacgtgct 133
Db 80 gctagaatccaatactaccctgtgcgcctgctcgtctgaacgtctccaataccaacgtgct 139
Qy 134 ctaaggttaacttcgatccgatccgatccgaagaatccaatccaatccagctgtccaatccgaatcct 193
Db 140 ctaaggttaacttcgatccgatccgatccgaagaatccaatccaatccagctgtccaatccgaatcct 199
Qy 194 ccaaaatcgaagtatactctgaagaatgctatcgtatatacaactctatgtacgaaaaaactctc 253
Db 200 ccaaatcgaagtatactctgaagaatgctatcgtatatacaactctatgtacgaaaaaactctc 259
Qy 254 ccactctcttcgtgactcgtatccgtatccgaataacttaactccaactctctcgtgaacaaatgat 313
Db 260 ccactctcttcgtgactcgtatccgtatccgaataacttaactccaactctctcgtgaacaaatgat 319
Qy 314 acaccatcaacaacgtgacatggaanaaactctcgtctggaagaatctctgaactacgcgtg 373
Db 320 acaccatcaacaacgtgacatggaanaaactctcgtctggaagaatctctgaactacgcgtg 379
Qy 374 aaatcatctggaactctgcgaagacactcgaagaatccaacagcgtgtgtatccaataact 433
Db 380 aaatcatctggaactctgcgaagacactcgaagaatccaacagcgtgtgtatccaataact 439
Qy 434 ctgagatgaatacaactctctgactatcatcaactcgtcgtgactctctgaacacccaaca 493
Db 440 ctgagatgaatacaactctctgactatcatcaactcgtcgtgactctctgaacacccaaca 499
Qy 494 atcgtctgaataactccaanaactctacatcaacgcgcgtctgatacgaccagaacagcatct 553
Db 500 atcgtctgaataactccaanaactctacatcaacgcgcgtctgatacgaccagaacagcatct 559
Qy 554 ccaatctgggtatacatccaacgcctctctaataacataatgttcaactgtgacggtgtctgctg 613
Db 560 ccaatctgggtatacatccaacgcctctctaataacataatgttcaactgtgacggtgtctgctg 619
Qy 614 acaactcaccgctacatcttgatccaataacttcaatctgttcgaacaagaactggaacga 673
Db 620 acaactcaccgctacatcttgatccaataacttcaatctgttcgaacaagaactggaacga 679
Qy 674 aagaataccaagaactgtgatacgacaacacagttccaattctgtatctctgaagaactcttg 733
Db 680 aagaataccaagaactgtgatacgacaacacagttccaattctgtatctctgaagaactcttg 739
Qy 734 gtgactaccctgcagtaagacaacacccgtactacatgtctgaatctgtacgataccgaacaa 793
Db 740 gtgactaccctgcagtaagacaacacccgtactacatgtctgaatctgtacgataccgaacaa 799
Qy 794 agcttgagctgaacaatgtaggtatctcgcggtgttaatatgtacacttggaagaagcttcgcg 853
Db 800 agcttgagctgaacaatgtaggtatctcgcggtgttaatatgtacacttggaagaagcttcgcg 859
Qy 854 cgttatgactacaacatctacactcgtgaactctctcctgtacgctgtgtacaaatccaatca 913
Db 860 cgttatgactacaacatctacactcgtgaactctctcctgtacgctgtgtacaaatccaatca 919
Qy 914 tcaagaatactgcgtctgtgtaacaagaacaaatctgttcgcaacaatgtatcgtgtatataca 973
Db 920 tcaagaatactgcgtctgtgtaacaagaacaaatctgttcgcaacaatgtatcgtgtatataca 979
Qy 974 tcaatgtctgtagttaagaacaagaatactacgctctcgtctaccaaagtctctcagcgtgctg 1033

Db 911 tcaagaatacgcgtctgtgtaacaagacatatcgttcgcaaatgatacgtgtatata 970
 Qy 974 tcaatgttgaatgaacaacaagatacgcgtctgtgtaacaatgcttcagcgtgtg 1033
 Db 971 tcaatgttgaatgaacaacaagatacgcgtctgtgtaacaatgcttcagcgtgtg 1030
 Qy 1034 tagaagaatctgtctgtcgtcgtgaatccgcgaactgtgtaactgtctcagtagtg 1093
 Db 1031 tagaagaatctgtctgtcgtcgtgaatccgcgaactgtgtaactgtctcagtagtg 1090
 Qy 1094 taatgaataccaagaacgagcggtatcactaacaagaataagaatcgtcgcaggaca 1153
 Db 1091 taatgaataccaagaacgagcggtatcactaacaagaataagaatcgtcgcaggaca 1150
 Qy 1154 acaatgtaacgatalcggattcaccgtgttcaccacgttcaacaatalcgttaaatcgtg 1213
 Db 1151 acaatgtaacgatalcggattcaccgtgttcaccacgttcaacaatalcgttaaatcgtg 1210
 Qy 1214 ttgcttccaactcgtgtacaaatcgtcagatcgaacgttcctcgtcagctcgtgtgtctt 1273
 Db 1211 ttgcttccaactcgtgtacaaatcgtcagatcgaacgttcctcgtcagctcgtgtgtctt 1270
 Qy 1274 gggagttcaccggttgatgagcgttggtggaacgttcgcgtgtgaa 1320
 Db 1271 gggagttcaccggttgatgagcgttggtggaacgttcgcgtgtgaa 1317

RESULT 5
 AAV30571
 ID AAV30571 standard; DNA; 1330 BP.
 XX

AAV30571;

07-DEC-1998 (first entry)

Clostridium botulinum toxin A fragment C gene in PalterBot.

Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen; botulism; ds.

Clostridium botulinum serotype A.

Key Location/Qualifiers

FT CDS 1..1317

FT misc_difference 1..6

FT /*tag= a

FT /*tag= b

FT /*note= "PALTER vector-derived nucleotides

FT (encode Met-Ala)"

PN W09808540-A1.

PD 05-MAR-1998.

PE 28-AUG-1997; 97WO-US15394.

PR 28-AUG-1996; 96US-0704159.

PA (OPHI-) OPHIDIAN PHARM INC.

PI Thalley BS, Williams JA;

DR WPI: 1998-230234/20.

DR P-PSDB; AAM68389.

XX

XX

XX

CC

CC C-fragment proteins have been produced in Escherichia coli as
 CC fusion proteins with either maltose binding protein or
 CC Clostridium difficile type A toxin (see AAM68387). The invention
 CC relates to recombinant proteins derived from C. botulinum toxins.
 CC Methods are provided which allow for the isolation of soluble
 CC recombinant proteins free of significant endotoxin contamination.
 CC Preferred hosts for production of recombinant toxin proteins are E. coli,
 CC insect cells and yeast cells. The recombinant toxin proteins are
 CC used as immunogens for the production of vaccines and antitoxins
 CC that are useful in the treatment of humans and animals at risk of
 CC intoxication with clostridial toxin.
 XX

Sequence 1330 BP; 400 A; 339 C; 246 G; 345 T; 0 other;

Query Match 98.2%; Score 1302.2; DB 19; Length 1330;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 ttgctctacctcactgaatacatcaagaacatcatacattccatcctcgtgaactcgc 73
 Db 11 ttgctctacctcactgaatacatcaagaacatcatacattccatcctcgtgaactcgc 70
 Qy 74 gctacgaatccaatccacccgtatcgcgtctgtcgaactcgttcgaatcgaatcgtgt 133
 Db 71 gctacgaatccaatccacccgtatcgcgtctgtcgaactcgttcgaatcgaatcgtgt 130
 Qy 134 cttaaatcactcgtatccgacatcgaagaatcgaatccagctgttcaatcgtgaactcgc 193
 Db 131 cttaaatcactcgtatccgacatcgaagaatcgaatccagctgttcaatcgtgaactcgc 190
 Qy 194 ccaaaatcgaagtatccctcgaagaatcgtatcgtatcaacaactcgtatcgaagaactcgc 253
 Db 191 ccaaaatcgaagtatccctcgaagaatcgtatcgtatcaacaactcgtatcgaagaactcgc 250
 Qy 254 ccaactcctcgtgacccgtatccggaatcactcaacccatcctcgtgaagaactcgc 313
 Db 251 ccaactcctcgtgacccgtatccggaatcactcaacccatcctcgtgaagaactcgc 310
 Qy 314 acacatcatcaactcgtatcgaagaacatcgttctggaagaatcgtatcgtgaactcgc 373
 Db 311 acacatcatcaactcgtatcgaagaacatcgttctggaagaatcgtatcgtgaactcgc 370
 Qy 374 aaatcatcgtgactcgtcgaagaactcgaagaatcgaagaactcgtatcgtgaactcgc 433
 Db 371 aaatcatcgtgactcgtcgaagaactcgaagaatcgaagaactcgtatcgtgaactcgc 430
 Qy 434 ctcaagatgataacatcgtatcgtatcgaagaatcgtatcgtatcgaagaactcgc 493
 Db 431 ctcaagatgataacatcgtatcgtatcgaagaatcgtatcgtatcgaagaactcgc 490
 Qy 494 atcgtctgaataactccaataatcacaatcgaagcgtcgtatcgaagaactcgc 553
 Db 491 atcgtctgaataactccaataatcacaatcgaagcgtcgtatcgaagaactcgc 550
 Qy 554 ccaatcgtggaatacgaactcgtatcgaagaactcgaagaactcgtatcgaagaactcgc 613
 Db 551 ccaatcgtggaatacgaactcgtatcgaagaactcgaagaactcgtatcgaagaactcgc 610
 Qy 614 acactcaccgtcactcgtatcgaagaactcgaagaactcgaagaactcgtatcgaagaactcgc 673
 Db 611 acactcaccgtcactcgtatcgaagaactcgaagaactcgaagaactcgtatcgaagaactcgc 670
 Qy 674 aagaatcaagaactcgtatcgaagaactcgaagaactcgaagaactcgtatcgaagaactcgc 733
 Db 671 aagaatcaagaactcgtatcgaagaactcgaagaactcgaagaactcgtatcgaagaactcgc 730
 Qy 734 gtgactactcgtggaatacgaagaactcgtatcgaagaactcgtatcgaagaactcgc 793
 Db 731 gtgactactcgtggaatacgaagaactcgtatcgaagaactcgtatcgaagaactcgc 790
 Qy 794 aggtgactcaagaatcgtatcgtatcgaagaactcgtatcgaagaactcgtatcgaagaactcgc 853

This is the DNA sequence of the Clostridium botulinum serotype A
 toxin C-fragment gene contained in plasmid PalterBot. Recombinant

```

Db 791 acgttgacatcaaatgtagtalcgcggtttacatgatactgaaggtccgctggtt 850
QY 854 ctgttatgactaacaaacttactgaacttctccctgtaccgtgtaccatca 913
Db 851 ctgttatgactaacaaacttactgaacttctccctgtaccgtgtaccatca 910
QY 914 tcaagaatacgcgtctgttaacaagaacatcgttcgcgaacatgacgtgtatca 973
Db 911 tcaagaatacgcgtctgtgttaacaagaacatcgttcgcgaacatgacgtgtatca 970
QY 974 tcaatgttgtatgaagaacaaatacgcgtctgtgtaccatgcttcaagctgtg 1033
Db 971 tcaatgttgtatgaagaacaaatacgcgtctgtgtaccatgcttcaagctgtg 1030
QY 1034 tagaagaatcttctgtctgtgtgaataccgcgagctgtgtatctgtctgaagtg 1093
Db 1031 tagaagaatcttctgtctgtgtgaataccgcgagctgtgtatctgtctgaagtg 1090
QY 1094 taatgaataccaagaacagcaggtatcactaacaatagtcaaatctgcagaga 1153
Db 1091 taatgaataccaagaacagcaggtatcactaacaatagtcaaatctgcagaga 1150
QY 1154 acaatgtgaacgatacgcgttcacgttccaccagttcaacaatacgtctaactg 1213
Db 1151 acaatgtgaacgatacgcgttccaccagttccaccagttcaacaatacgtctaactg 1210
QY 1214 ttgtctccaactgttacaatgtagatgcgaacttctctcgcgactctgtgtctt 1273
Db 1211 ttgtctccaactgttacaatgtagatgcgaacttctctcgcgactctgtgtctt 1270
QY 1274 gggagttcatcccggttgatgaaggttgagggtgaacgtccgctgtaa 1320
Db 1271 gggagttcatcccggttgatgaaggttgagggtgaacgtccgctgtaa 1317

RESULT 6
AAZ87212
ID AAZ87212 standard; DNA; 1338 BP.
AC AAZ87212;
XX
XX 08-MAY-2000 (first entry)
DE DNA encoding synthetic BONT serotype A (BONTA) Hc fragment.
XX
XX Botulinum neurotoxin; heavy chain: BONT; serotype A;
KM C-terminal fragment; Venezuelan equine encephalitis virus replicon;
KM VEE; botulinum; vaccine; diagnosis; drug screening; ds.
XX
XX Clostridium botulinum.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
XX CDS 9..1325
XX /*tag= "a
XX /product= "Synthetic botulinum neurotoxin serotype A
XX (BONTA) heavy chain C-terminal fragment (Hc)"
XX
XX MO200002524-A2.
XX
XX 20-JAN-2000.
XX
XX 09-JUL-1999; 99WO-US15570.
XX
XX 10-JUL-1998; 98US-0092416.
XX 12-MAY-1999; 99US-0133870.
XX
XX (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
XX Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
XX WPI; 2000-160827/14.
XX
XX P-PSDB; AAY77134.
XX

```

```

XX
XX Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
XX toxin serotypes A-G, is used for inducing an immune response against
XX botulinum -
XX
XX PS Disclosure: Page 54; 54pp; English.
XX
XX The invention relates to novel vaccines that induce a protective immune
XX response against botulinum neurotoxin (BONT) serotypes A, B, C, D, E, F
XX and G (BONTA-BONTG). The vaccine of the invention is novel recombinant
XX DNA construct comprising a C-terminal heavy chain fragment (Hc) from BONT
XX fragment comprising a C-terminal heavy chain fragment (Hc) from BONT
XX serotypes A-G. In preferred embodiments of the invention, the vector is
XX a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of
XX this vector results in the production of large amounts of a protein
XX encoded by a sequence cloned into the replicon. The constructs are used
XX to produce vaccines against botulinum. The proteins can also be used as
XX diagnostic tools for the diagnosis of botulinum. The transformed host
XX cells can be used to analyse the effectiveness of drugs and agents which
XX inhibit toxin effects. The vaccine currently used against botulinum is
XX dangerous and expensive to produce, and contains formalin, which is very
XX painful for the recipient. Also, the vaccine is incomplete, in that only
XX 5 of the 7 serotypes are represented in the formulation. The novel
XX vaccine of overcomes these problems, as it is easily purified, and
XX available in large quantities. It is also expressed in the lymph nodes
XX for a better immune response. Sequences AAZ87212-287217 represent
XX synthetic DNA sequences encoding BONT Hc fragments used in the present
XX invention. These were optimised for codon usage for expression in yeast.
XX
XX Sequence 1338 BP; 401 A; 342 C; 249 G; 346 T; 0 other:

```

```

Query Match 98.2%; Score 1302.2; DB 21; Length 1338;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 14 tggcctctactctactgatacatcaagaacatcaatacctccactcgtgaactgc 73
Db 19 tggctgtcactctactcgtgatacatcaagaacatcaatacctccactcgtgaactgc 78
QY 74 gctacgaatccaatcaactgtgacgtctgtcgtgaagcttccaaatacaactcgtgt 133
Db 79 gctacgaatccaatcaactgtgacgtctgtcgtgaagcttccaaatacaactcgtgt 138
QY 134 ctaagtttaactcgtatccgatacgaagaacatcagaatcgttcaactcgtgaacttc 193
Db 139 ctaagtttaactcgtatccgatacgaagaacatcagaatcgttcaactcgtgaacttc 198
QY 194 ccaaatcgaagttaactcgtgaagaatgtatcgtatcaactctatgtacgaactctc 253
Db 199 ccaaatcgaagttaactcgtgaagaatgtatcgtatcaactctatgtacgaactctc 258
QY 254 ccaactcctctctgtatcgtatccgaaatcaactcaactcctctctgaacatgaat 313
Db 259 ccaactcctctctgtatcgtatccgaaatcaactcaactcctctctgaacatgaat 318
QY 314 acaccatacgaactcgtgatacgaagaacatcgtgttggaagaatcctcgtgaactcgtgt 373
Db 319 acaccatacgaactcgtgatacgaagaacatcgtgttggaagaatcctcgtgaactcgtgt 378
QY 374 aatcatctggaactcgtgacgacactcagaagaatcaaaacggttgttatccaataact 433
Db 379 aatcatctggaactcgtgacgacactcagaagaatcaaaacggttgttatccaataact 438
QY 434 ctcaatgatcaacatctctgatacatcaatcgtctgtatcttgattaccatccaaca 493
Db 439 ctcaatgatcaacatctctgatacatcaatcgtctgtatcttgattaccatccaaca 498
QY 494 atcgtcgaataacccaacatcctacatcaacgagcgctgtgtgacgaagaacgatct 553
Db 499 atcgtcgaataacccaacatcctacatcaacgagcgctgtgtgacgaagaacgatct 558
QY 554 ccaatctgggttaacatccacgcttctaatacatcatgttcaactggaaggttgtcgtg 613

```



```

Db      559 ccaatctggtaacatcccgctctctaatacatcatgttcaaacctggacggtgttcgtg 618
Oy      614 acactaccgtacatcatgatcaaaactcaatctgttctgcgaagaactgaagaaa 673
Db      619 acactaccgtacatcatgatcaaaactcaatctgttctgcgaagaactgaagaaa 678
Oy      674 aagaatacaagaacctgtacgaacaacagtcacaattctgtatctctgaaagactctggg 733
Db      679 aagaatacaagaacctgtacgaacaacagtcacaattctgtatctctgaaagactctggg 738
Oy      734 gtagactactgtacgtacgaacaacagtcacaattctgtatctctgaaagactctggg 733
Db      739 gtagactactgtacgtacgaacaacagtcacaattctgtatctctgaaagactctggg 738
Oy      794 acgttgaactcaacaatgttaggtatccgcggtatcatgttaccggaagctccgctgggt 853
Db      799 acgttgaactcaacaatgttaggtatccgcggtatcatgttaccggaagctccgctgggt 858
Oy      854 ctgttatgactacaacatctacactgtacactctccctgtacacgttggtaaccaaatcatca 913
Db      859 ctgttatgactacaacatctacactgtacactctccctgtacacgttggtaaccaaatcatca 918
Oy      914 tcaagaataacgctctgtgtataaagaacaatactgttgcgaacaatgatcgtgtatata 973
Db      919 tcaagaataacgctctgtgtataaagaacaatactgttgcgaacaatgatcgtgtatata 978
Oy      974 tcaatcttgaatgaagaacaagaatacaccgtctgtgctacaaatgtcttcagacgtggtg 1033
Db      979 tcaatcttgaatgaagaacaagaatacaccgtctgtgctacaaatgtcttcagacgtggtg 1038
Oy      1034 tgaagaagaactctgtctgtctgtgaataaccgcggaactgtgttaactctgtcctcaggtg 1093
Db      1039 tgaagaagaactctgtctgtctgtgaataaccgcggaactgtgttaactctgtcctcaggtg 1098
Oy      1094 tatgtaatccaaagaacgacccgggtgtatatacctaacaatggaatctgtcagaaga 1153
Db      1099 tatgtaatccaaagaacgacccgggtgtatatacctaacaatggaatctgtcagaaga 1158
Oy      1154 acaatgttaacgatatccgtttcatcggttccacagttccaacaatcgttaactg 1213
Db      1159 acaatgttaacgatatccgtttcatcggttccacagttccaacaatcgttaactg 1218
Oy      1214 ttgcttccaaactgtgtacaactcgttcagatcgaacgttctctctcgaactctgtgtctct 1273
Db      1219 ttgcttccaaactgtgtacaactcgttcagatcgaacgttctctctcgaactctgtgtctct 1278
Oy      1274 gggagttcatcccggttgaatgaacgttgggtggaacgtccgcgtgtaa 1320
Db      1279 gggagttcatcccggttgaatgaacgttgggtggaacgtccgcgtgtaa 1325

RESULT
7
AAV30576 ID AAV30576 standard; DNA; 1351 BP.
XX
AC AAV30576;
XX
DT 07-DEC-1998 (first entry)
XX
DE Clostridium botulinum toxin A fragment C gene in pHisBotA(syn).
XX
KW AntiToxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
KW botulism; ds.
XX
OS Clostridium botulinum serotype A.
XX
XX Synthetic.
XX
Key Location/Qualifiers
FH CDS 1..1338
FT /*tag= a
PN WO9808540-A1.

```

```

XX
PD 05-MAR-1998.
XX
PF 28-AUG-1997; 97WO-US15394.
XX
PR 28-AUG-1996; 96US-0704159.
XX
PA (OPHI-) OPHIDIAN PHARM INC.
XX
PI Thalley BS, Williams JA.
XX
DR WPI: 1998-230234/20.
DR P-PSDB; AAW68391.
XX
PT Host cell containing recombinant expression vector encoding
PT Clostridium botulinum type B or E toxin - useful to treat humans
PT and other animals at risk of intoxication with clostridial toxin
XX
PS Example 29; Page 279-281; 428pp; English.
XX
CC This is the DNA sequence of the Clostridium botulinum serotype A
CC toxin C fragment gene contained in plasmid pHisBotA(syn). The
CC encoded toxin A polypeptide (see AAW68391) has a histidine-tagged
CC N-terminal extension. The vector was used to express native
CC (i.e. non-fusion) soluble C fragment in Escherichia coli host
CC cells. The invention relates to recombinant proteins derived from
CC C. botulinum toxins. Methods are provided which allow for the
CC isolation of soluble recombinant proteins free of significant
CC endotoxin contamination. Preferred hosts for production of
CC recombinant proteins are E. coli, insect cells and yeast cells.
CC The recombinant toxins are used as immunogens for the production
CC of vaccines and antitoxins that are useful in the treatment of
CC humans and animals at risk of intoxication with clostridial toxin.
XX
SQ Sequence 1351 BP; 407 A; 348 C; 247 G; 349 T; 0 other:

Query Match 98.2%; Score 1302.2; DB 19; Length 1351;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 14 ttgccccttaactcactgaatatacataagaacatcataactcatctgaactcgc 73
Db 32 ttgtctgtattactcttaactgaatatacataagaacatcataactcatctgaactcgc 91
Oy 74 gctacgaatccaatcacactgatcgaactgtctcgtctcgaagttccaaatcaacatcggt 133
Db 92 gctacgaatccaatcacactgatcgaactgtctcgtctcgaagttccaaatcaacatcggt 151
Oy 134 ctaaaagttaaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 193
Db 152 ctaaaagttaaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 211
Oy 194 ccaaaatcgaagttaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 253
Db 212 ccaaaatcgaagttaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 271
Oy 254 ccaactcctcttgatccgatccgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 313
Db 272 ccaactcctcttgatccgatccgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 331
Oy 314 acacacatcatcaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 373
Db 332 acacacatcatcaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 391
Oy 374 aaatccttgactctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 433
Db 392 aaatccttgactctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 451
Oy 434 ctcaatgatcaacatctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 493
Db 452 ctcaatgatcaacatctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 511

```


QY 494 atcgtctgaataactccaataactatcaatcaacggcgctctgatacgaacagaaccgactc 553
|||||
Db 512 atcgtctgaataactccaataactatcaatcaacggcgctctgatacgaacagaaccgactc 571
QY 554 ccaatctgggtgaataactccaactcttaataatcaatggttcaactggaggtgtcggt 613
|||||
Db 572 ccaatctgggtgaataactccaactcttaataatcaatggttcaactggaggtgtcggt 631
QY 614 acaatcaacgctatcaatctgggttaataatcaatctgtctgacagaagaagaacaggagaa 673
|||||
Db 632 acaatcaacgctatcaatctgggttaataatcaatctgtctgacagaagaagaacaggagaa 691
QY 674 aagaatacaagaactgtgacagaacaaccgagtcgaattccgtgatacctgtaagaactctgg 733
|||||
Db 692 aagaatacaagaactgtgacagaacaaccgagtcgaattccgtgatacctgtaagaactctgg 751
QY 734 gtgatactctgacgtgaagaacaaccgatactatactgataatctgatacgcgaacaat 793
|||||
Db 752 gtgatactctgacgtgaagaacaaccgatactatactgataatctgatacgcgaacaat 811
QY 794 acaatgaatgaataatgttaagttatccggttataatctgataatgaagaagtcgggtgt 853
|||||
Db 812 acaatgaatgaataatgttaagttatccggttataatctgataatgaagaagtcgggtgt 871
QY 854 ctgtatgataccaacaactcaactgaactcttccctgtacggtgataccaataatcatca 913
|||||
Db 872 ctgtatgataccaacaactcaactgaactcttccctgtacggtgataccaataatcatca 931
QY 914 tcaagaataacgctctgtgatacaagaacaataatctgtccgaacaatgatactatata 973
|||||
Db 932 tcaagaataacgctctgtgatacaagaacaataatctgtccgaacaatgatactatata 991
QY 974 tcaatgtgtgataagaacaagaacgctctgtgataccaactcttccgaaggtgtgt 1033
|||||
Db 992 tcaatgtgtgataagaacaagaacgctctgtgataccaactcttccgaaggtgtgt 1051
QY 1034 tagaanaagatctgtctgtctgtgaatcccgagactgtgtaatctgtctcagtagtgt 1093
|||||
Db 1052 tagaanaagatctgtctgtctgtgaatcccgagactgtgtaatctgtctcagtagtgt 1111
QY 1094 taagaataccaagaacgacgaggtatcaactataaataatgataatctgtccgagaa 1153
|||||
Db 1112 taagaataccaagaacgacgaggtatcaactataaataatgataatctgtccgagaa 1171
QY 1154 acaatgtgaacgatactcggttcaatcggttccacaggttcaacaatctgtcaactgtg 1213
|||||
Db 1172 acaatgtgaacgatactcggttcaatcggttccacaggttcaacaatctgtcaactgtg 1231
QY 1214 ttgcttcaactgtgatacaatcggtcagatcgaaagttcctctcgaactctgggtgtgtct 1273
|||||
Db 1232 ttgcttcaactgtgatacaatcggtcagatcgaaagttcctctcgaactctgggtgtgtct 1291
QY 1274 gggaggtatcatcccggtgatacaggttggggtgaacgtccgctgttaa 1320
|||||
Db 1292 gggaggtatcatcccggtgatacaggttggggtgaacgtccgctgttaa 1338

RESULT 8

AAT29246
ID AAT29246 standard; DNA; 1402 BP.

XX AAT29246;

XX 07-JUL-1996 (first entry)

XX Type A neurotoxin C fragment-polyhistidine tag gene fusion.

XX Toxin; neurotoxin; fusion protein; antitoxin; vaccine; immunogen;

XX Clostridium botulinum; polyhistidine; vector; PERTHisa; phisbot; ds.

OS Synthetic.
XX
FH Key Location/Qualifiers

FT CDS 1..1317
FT /*tag= a
FT /product= phisBot fusion protein
PN W09612802-A1.
PD 02-MAY-1996.
XX
XX 23-OCT-1995; 95WO-US13737.
XX
XX 07-JUN-1995; 95US-0480604.
PR 24-OCT-1994; 94US-0329154.
PR 16-MAR-1995; 95US-0405496.
PR 14-APR-1995; 95US-0422711.
XX
PA (OPHI-) OPHIDIAN PHARM INC.
XX
XX Frica JR, Kink JA, Padhye NV, Stafford DC, Thalley BS;
PI Williams JA;
XX
DR MPI: 1996-230603/23.
XX P-PSDB: AAR95009.
XX
XX Fusion proteins comprising non-toxin protein and part of toxin
XX useful to form anti-toxins against Clostridium botulinum type A, and
XX C. difficile type toxins, and to treat C. difficile intoxication,
XX partic. diarrhoea
XX
XX Example 24; Page 340-342; 434pp; English.
XX
XX A nucleotide sequence (AAT29246) present in vector PERTHisa encodes
XX the phisBot fusion protein (AAR95009) comprising a polyhistidine
XX affinity tag and fragment C (see also AAR95008) of the Clostridium
XX botulinum type A neurotoxin. The phisBot protein was expressed
XX in Escherichia coli as a soluble protein and was purified by
XX metal chelate affinity chromatography to obtain a product free
XX of endotoxin contamination that may be useful as an immunogen
XX in vaccine compns.
XX
SQ Sequence 1402 BP; 420 A; 360 C; 260 G; 362 T; 0 other;

Query Match 98.2%; Score 1302.2; DB 17; Length 1402;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 tggctctaccttcaactgataatcaatcaagaataatcaatcaatcttccatctgaaactgc 73
|||
Db 83 tggctctaccttcaactgataatcaatcaagaataatcaatcaatcttccatctgaaactgc 142
QY 74 gctacgaatccaataccctgatacgtctctcgtacgattcccaaatccaatcggtt 133
|||||
Db 143 gctacgaatccaataccctgatacgtctctcgtacgattcccaaatccaatcggtt 202
QY 134 cttaagtttaacttcgataccgatacgaagaatccaatccaatggttcaatctgaaactct 193
|||||
Db 203 cttaagtttaacttcgataccgatacgaagaatccaatccaatggttcaatctgaaactct 262
QY 194 ccaaaatcgaaagttaactcgaagaatgatacgtatatacaaacctatgatacgaagaattct 253
|||||
Db 263 ccaaaatcgaaagttaactcgaagaatgatacgtatatacaaacctatgatacgaagaattct 322
QY 254 ccacactccttggatcgatcccgaaataacttaactcaactcctctcgtgaacaatgaaat 313
|||||
Db 323 ccacactccttggatcgatcccgaaataacttaactcaactcctctcgtgaacaatgaaat 382
QY 314 acacatcatcaactgatacgtgaagaacaactctggttgaagaatctcgtgaactacggtgt 373
|||||
Db 383 acacatcatcaactgatacgtgaagaacaactctggttgaagaatctcgtgaactacggtgt 442
QY 374 aaatactctgacacttgacagacactcaagaataatcaaacagcgtgtgtatcaataact 433
|||||
Db 443 aaatactctgacacttgacagacactcaagaataatcaaacagcgtgtgtatcaataact 502

Db 383 acaccatcaactcgtatggaacaaatctggttggaagatctctctgaactcagtg 442
 QY 374 aaatcatctgactctgcagagacactcaggaatacaacagcggttggtatcaaatc 433
 Db 443 aaatcatctgactctgcagagacactcaggaatacaacagcggttggtatcaaatc 502
 QY 434 ctcaagatcaaatctctgactacatacactcgtctgactctctgactacatac 493
 Db 503 ctcaagatcaaatctctgactacatacactcgtctgactctctgactacatac 562
 QY 494 atgctcgtataactccaataatctacatacagcgctctgatacagacaaccgact 553
 Db 563 atcgtctgataatccaataatctacatacagcgctctgatacagacaaccgact 622
 QY 554 ccaatcgtggttaacatcacgcgtcttaataacatcagttcaaatctgagcgttgctg 613
 Db 623 ccaatcgtggttaacatcacgcgtcttaataacatcagttcaaatctgagcgttgctg 682
 QY 614 acaatcacgcgtacatctgatacaataactcaactctgcttcgacaagaactgagaa 673
 Db 683 acaatcacgcgtacatctgatacaataactcaactctgcttcgacaagaactgagaa 742
 QY 674 aagaatacaagaacgcttgaagacaacacagtcacattctgatacctgaaagactctggg 733
 Db 743 aagaatacaagaacgcttgaagacaacacagtcacattctgatacctgaaagactctggg 802
 QY 734 gtgactacgctgaactcagacaacacgctactacatcgtctgaactctgatacctcagacaat 793
 Db 803 gtgactacgctgaactcagacaacacgctactacatcgtctgaactctgatacctcagacaat 862
 QY 794 acgttgaactcagacaacacgctactacatcgtctgaactctgatacctcagacaat 853
 Db 863 acgttgaactcagacaacacgctactacatcgtctgaactctgatacctcagacaat 922
 QY 854 ctgcttgaactcagacaacacgctactacatcgtctgaactctgatacctcagacaat 913
 Db 923 ctgcttgaactcagacaacacgctactacatcgtctgaactctgatacctcagacaat 982
 QY 914 tcaagaatacgcgtctcgtgaagacaacatcgtcttcgacaacatcgtctgatac 973
 Db 983 tcaagaatacgcgtctcgtgaagacaacatcgtcttcgacaacatcgtctgatac 1042
 QY 974 tcaatgttgaactgaagaacaagaatacgcgtctcgtgaacatcgtctcagagctg 1033
 Db 1043 tcaatgttgaactgaagaacaagaatacgcgtctcgtgaacatcgtctcagagctg 1102
 QY 1034 tagaagaatctctgctcgtctggaatacgcgacgcttggtatcctgctcaggtagt 1093
 Db 1103 tagaagaatctctgctcgtctggaatacgcgacgcttggtatcctgctcaggtagt 1162
 QY 1094 taatgaataccaagaacagcaggtatcactaacaatgtaaaatgtaattgcagagca 1153
 Db 1163 taatgaataccaagaacagcaggtatcactaacaatgtaaaatgtaattgcagagca 1222
 QY 1154 acaatgttaacgataatcgtttcactcgtttccaccagttcaacaatacgtctaaacg 1213
 Db 1223 acaatgttaacgataatcgtttcactcgtttccaccagttcaacaatacgtctaaacg 1282
 QY 1214 ttgcttccaactcgttaacaatcgtcagatcgaagcttcctcctcgaactctgggtgct 1273
 Db 1283 ttgcttccaactcgttaacaatcgtcagatcgaagcttcctcctcgaactctgggtgct 1342
 QY 1274 gggagttcattccgcgttgatagcaggttggtggaacgtcgcgtgtaa 1320
 Db 1343 gggagttcattccgcgttgatagcaggttggtggaacgtcgcgtgtaa 1389

RESULT 10
 AAA54588
 ID AAA54588 standard; DNA: 1332 BP.
 AC AAA54588;
 XX

DT 11-APR-2001 (first entry)
 XX
 DE Sequence encoding botulism toxin C fragment (serotype A).
 XX
 KW Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 KW infection; ds.
 OS
 OS Clostridium botulinum.
 XX
 PN W020067700-A2.
 PD 16-NOV-2000.
 PD 12-MAY-2000; 2000WO-0512890.
 XX
 PR 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.
 XX
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 XX
 DR WPI: 2001-016048/02.
 XX
 PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism
 XX
 PS
 PS
 PS Example 7; Page 37; 73pp; English.
 CC Botulism neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dimer
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.
 XX
 SQ Sequence 1332 BP; 401 A; 342 C; 244 G; 344 T; 1 other;

Query Match 94.1%; Score 1247.8; DB 22; Length 1332;
 Best Local Similarity 99.2%; Pred. NO. 0;
 Matches 1296; Conservative 1; Mismatches 4; Indels 6; Gaps 4;

QY 14 tggcctcactcactcagtaatacaagaataatcaatacctccatcctgaacctgc 73
 Db 19 tgcgtctcactcactcagtaatacaagaataatcaatacctccatcctgaacctgc 78
 QY 74 gctacgaatccaatcactgtagcactgctcgtcgtacgcttccaataatcaatcaggt 133
 Db 79 gctacgaatccaatcactgtagcactgctcgtcgtacgcttccaataatcaatcaggt 138
 QY 134 cttaagtttaactcgtatcgcagcaagaatcaatcaggttggttaactcgtgaactct 193
 Db 139 cttaacttaactcgtatcgcagcaagaatcaatcaggttggttaactcgtgaactct 198

```

OY 194 ccaaaatcgagttatctctgaagaatgctatcgtatatacaactatgtagcgaagaactct 253
    |||
DB 199 ccaaaatcgagttatctctgaagaatgctatcgtatatacaactatgtagcgaagaactct 258
OY 254 ccacccctctgagtcggtatcccgaaatctcaactccatctctctcgaagaatgtagt 313
    |||
DB 259 ccacccctctgagtcggtatcccgaaatctcaactccatctctctcgaagaatgtagt 317
OY 314 acacacatcaactgtagcgaagaagaatctggtgtagaagtagtctcgaactgtagt 373
    |||
DB 318 acacacatcaactgtagcgaagaagaatctggtgtagaagtagtctcgaactgtagt 377
OY 374 aaatcatctgagtcgtagcgaagaactcgaagaatacaacggtgtgttatacttaact 433
    |||
DB 378 aaatcatctgagtcgtagcgaagaactcgaagaatacaacggtgtgttatacttaact 437
OY 434 ctccagatgtagcaactctgtagtatactcaatcgctgtagtcttcgttaccatccaaca 493
    |||
DB 438 ctccagatgtagcaactctgtagtatactcaatcgctgtagtcttcgttaccatccaaca 497
OY 494 atcgctctgtagtaactccaaaatctacatcaacggtgtagtctgtagcgaagaacgtagt 553
    |||
DB 498 atcgctctgtagtaactccaaaatctacatcaacggtgtagtctgtagcgaagaacgtagt 556
OY 554 ccaatctggtgtagtaactccacgcttcttaataacatcatgttcaactgtagcgtgtgtct 613
    |||
DB 557 ccaatctggtgtagtaactccacgcttcttaataacatcatgttcaactgtagcgtgtgtct 615
OY 614 acacacacgctgtagtctgtagtatactcaatctggtgtagcgaagaactgtagcgaaga 673
    |||
DB 616 acacacacgctgtagtctgtagtatactcaatctggtgtagcgaagaactgtagcgaaga 675
OY 674 aagaatcaagaactgtagcgaagaacagttcaatctggtgtagcgaagaactctgtagt 733
    |||
DB 676 aagaatcaagaactgtagcgaagaacagttcaatctggtgtagcgaagaactctgtagt 735
OY 734 gtagtactcgtgtagcgaagaacagttactatctgtagtctgtagcgtctcgaagaact 793
    |||
DB 736 gtagtactcgtgtagcgaagaacagttactatctgtagtctgtagcgtctcgaagaact 795
OY 794 acgttgaagtcgaagaatgtagtatactcgtgtagtactgtagcgaagaactcgtgtagt 853
    |||
DB 796 acgttgaagtcgaagaatgtagtatactcgtgtagtactgtagcgaagaactcgtgtagt 855
OY 854 cgtttagtactcgaagaatctacactgtagtctcgtgtagtactcgtgtagtactcaatca 913
    |||
DB 856 cgtttagtactcgaagaatctacactgtagtctcgtgtagtactcgtgtagtactcaatca 915
OY 914 tcaagaataagcgtctgtagtatactcgtgtagtactcgtgtagtactcgtgtagtactca 973
    |||
DB 916 tcaagaataagcgtctgtagtatactcgtgtagtactcgtgtagtactcgtgtagtactca 972
OY 974 tcaatggtgtagttagaagaagaatatactgtagtactcgtgtagtactcgtgtagtact 1033
    |||
DB 973 tcaatggtgtagttagaagaagaatatactgtagtactcgtgtagtactcgtgtagtact 1032
OY 1034 tagaagaagaatctgtagtactcgtgtagtactcgtgtagtactcgtgtagtactcgtgtagt 1093
    |||
DB 1033 tagaagaagaatctgtagtactcgtgtagtactcgtgtagtactcgtgtagtactcgtgtagt 1092
OY 1094 taatgaataatcgaagaagcaggttagtactcaactaagaagaatgaatgtagttagcagaga 1153
    |||
DB 1093 taatgaataatcgaagaagcaggttagtactcaactaagaagaatgaatgtagttagcagaga 1152
OY 1154 acaatgtagaagatatactgtagtactcgtgtagtactcgtgtagtactcgtgtagtact 1213
    |||
DB 1153 acaatgtagaagatatactgtagtactcgtgtagtactcgtgtagtactcgtgtagtact 1212
OY 1214 ttgcttcaactgtagtactcgtgtagtactcgtgtagtactcgtgtagtactcgtgtagtact 1273
    |||
DB 1213 ttgcttcaactgtagtactcgtgtagtactcgtgtagtactcgtgtagtactcgtgtagtact 1272
OY 1274 gggagttatcatcccggttagttagcgtgtgggtgtagaagtcgcgctgttaa 1320

```

```

DB 1273 gggagttatcatcccggttagttagcgtgtgggtgtagaagtcgcgctgttaa 1319
    |||
RESULT 11
AAV30575
ID AAV30575 standard; DNA; 1546 BP.
XX
AC AAV30575;
XX
DT 07-DEC-1998 (first entry)
XX
DE Clostridium botulinum toxin A fragment C gene in phisBoA.
XX
KW Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
XX botulism; ds.
XX
OS Clostridium botulinum serotype A.
XX
FH Key Location/Qualifiers
FT CDS 108..1496
FT /tag= a
XX
XX WO9808540-A1.
XX
XX 05-MAR-1998.
XX
XX 28-AUG-1997; 97WO-US15394.
XX
XX 28-AUG-1996; 96US-0704159.
XX
XX (OPHI-) OPHIDIAN PHARM INC.
XX
XX Thalley BS, Williams JA;
XX
XX WPI: 1998-230234/20.
XX
XX P-PSDB: AAW68390.
XX
XX Host cell containing recombinant expression vector encoding
XX Clostridium botulinum type B or E toxin - useful to treat humans
XX and other animals at risk of intoxication with clostridial toxin
XX
XX Example 28; Page 277-278; 428pp; English.
XX
XX This is the DNA sequence of the Clostridium botulinum serotype A
XX toxin C fragment gene contained in plasmid pHisBoA. The encoded
XX toxin A polypeptide (see AAW68390) has a histidine-tagged
XX N-terminal extension. The vector was used to express native
XX (i.e. non-fusion) soluble C fragment in Escherichia coli host
XX cells. The invention relates to recombinant proteins derived from
XX C. botulinum toxins. Methods are provided which allow for the
XX isolation of soluble recombinant proteins free of significant
XX endotoxin contamination. Preferred hosts for production of
XX recombinant proteins are E. coli, insect cells and yeast cells.
XX The recombinant toxins are used as immunogens for the production
XX of vaccines and antitoxins that are useful in the treatment of
XX humans and animals at risk of intoxication with clostridial toxin.
XX
XX Sequence 1546 BP; 629 A; 163 C; 256 G; 498 T; 0 other;

```

Query Match 54.8%; Score 727.2; DB 19; Length 1546;
Best Local Similarity 72.2%; Pred. No. 1.2e-194;
Matches 945; Conservative 0; Mismatches 363; Indels 0; Gaps 0;

```

OY 19 1ttacctcaatgtagaatacaagaacatcaataactccatccatcgtgaactggtctac 78
    |||
DB 195 1ttacctcaatgtagaatacaagaacatcaataactccatccatcgtgaactggtctac 254
OY 79 gatccaatcactgtagcgtgtctgtagtactcgtgtagtactcgtgtagtactcgtgtagt 138
    |||
DB 255 gatccaatcactgtagcgtgtctgtagtactcgtgtagtactcgtgtagtactcgtgtagt 314

```


Query Match	54.8%;	Score 727.2;	DB 21;	Length 4835;
Best Local Similarity	72.2%;	Pred. No. 2.1e-194;		
Matches	945;	Conservative	0;	Mismatches 363;
			Indels	0;
			Gaps	0;
QY	19	tctactctacatgaataactcaagaacatcatcaataactccatccttgaaactcgctac	78	
DB	2947	tctaaattactgaataataatgaataattatcaattactctataattgaattgaataat	3006	
QY	79	gaatcaaatccatgtaacacgtctcgcgtccacgtcccaaatccgaactcggtctaa	138	
DB	3007	gaagaatcaatcttaataataactatctatagatgcatcaaaaaataatattgtaata	3066	
QY	139	gttaactctgatacgcgtacgaacaagaatcagatccagctgttcaactcgtgaactc	198	
DB	3067	gttaaatcttgatccaataagaataaaatcaaatcaattatctaattagaagtaata	3126	
QY	199	atcgaagtatccctgaagaatgctatcgtatcaactcatgtatgagaactctccacc	258	
DB	3127	atggaagtaattttaaaaaatgctattgtatataatagatagatgaataatttagtact	3186	
QY	259	tcccttgatccgtatcccgaaataactcaactcatcctcctgacaatgataacac	318	
DB	3187	agcttttgataagaattccctctaagtatttaacagataagtcataataatgataata	3246	
QY	319	atcaataactgcatggaaaaaaatctcgtgtggaagatctctcgtgaactcgtgaact	378	
DB	3247	ataataaatctgatatggaataaataatcagaatggaagatcaacttaattatgtgaata	3306	
QY	379	atctggaactctgacggaactcaggaataccaacagcgtgtgattccaataactctcg	438	
DB	3307	atctggaactcttaagaatctcccggaataaacaagaagtagtcttctaataacagtcga	4366	
QY	439	atgatcaaatctctgatactacatcaatcgtcgtgatactctgtaaccaaccaaatctg	498	
DB	3367	atgatttaataatcagatatataataacagatggaattcttgtaactatcaactaataaga	3426	
QY	499	ctggaataactccaataatctcaatcaacgagcgtctgatacgcagaaacgcagatccaat	558	
DB	3427	ttaaaataactcttaaaattatataaaatggaagaataatagatccaataatccaat	3486	
QY	559	ctgggtaacatccacgctcttaataacatcatgattcaactgagcggtctgtcgtgacact	618	
DB	3487	ttaggttaataatcattcatgtaataataatgatttaaatatgaatggttgaagataca	3546	
QY	619	caacgctacatctggaatcaatactcaatctgttcgacaagaatgacgaaaaaaga	678	
DB	3547	catagatatatttgataaaataatttaactcttcttgataagaatataatgaaaaaga	3606	
QY	679	atcaaaagacctgtaacgaacaacagatccaattctgtatccctgaaagactctggggtagac	738	
DB	3607	atcaaaagatttatatgataatcaatcaaatctcagttattttaaagaacttttggggtagat	3666	
QY	739	taccctgagtagaacaacacgtaactacatgctgaatcgttagactcgcgaacaatcgct	798	
DB	3667	tatttaacaataatgaataaccatactatagtttaaatattatagatccaataataatgct	3726	
QY	799	gaacgtcaaacatgtagtatcccggtttacatgtaacgttcaaaagttcgcgtgtctgtt	858	
DB	3727	gattgataaataatgtagtatttagaggttatatgtatcttaaaagcgcttagagtagcga	3786	
QY	859	atgactacacaacatctcaactgaactctccctgtaacccgtgtagtaaccaaatcatcatcaag	918	
DB	3787	atgactacacaacatcttatttaaatccaagtttgtagaggggagcaaaaattattatacaaa	3846	
QY	919	aaataacgctctgtaagaagaacatcatcgttcgacaacatgatatcatcaat	978	
DB	3847	aaataacgctctcggaaataaagaataatctgttagaataaataatgtagtataataat	3906	
QY	979	gttgttagtaagaacaagaataacgctcgtgataccaatcgtctcgaacgtgtgtgagaa	1038	
DB	3907	gttagtagttaaaataaagaataatagtttagctactaactgcatcacaaagcagagcgtgagaa	3966	
QY	1039	aagatctgtctgtcctcgtgaataatcccgagcgttgtagtaactcgtctcagtagttagtaacgt	1098	

DB	3967	aaataactaagtcattagaataactcgtatgtaggaatactaaagtcagaagtagtaacgt	4026	
QY	1099	aaatccaagaagcagcagggtatcactaaacaatgcaaaatgatctgcagacaacaat	1158	
DB	4027	aagtcacaaaatgataaagaataaacaataaactgaacaaatgataatctacaagataat	4086	
QY	1159	ggttaacgataatcggtttccatcggtttccacaggttcaacaataatcgttaaatggttgc	1218	
DB	4087	gggaatgatataagctttatagatttcacgtatcagtttaataatagataactagtagca	4146	
QY	1219	tccaactggtacaaatcgttaagatgaacgttccctcgcgaactcgtgtgtctgtggag	1278	
DB	4147	agtaattggtataatagacaataatagaagaatcagtagaggaacttgggtgtcattgga	4206	
QY	1279	tccatccggtttagagcgttggtggaacgtccggtctgaagaattc	1326	
DB	4207	ttatctccgttagatgataatgtagtgagggaagaagccacgttaataatc	4254	
RESULT 13				
AAZ87220				
ID	AAZ87220	standard; DNA; 1317 BP.		
XX	XX	AAZ87220;		
AC	XX	08-MAY-2000 (first entry)		
XX	XX			
DE	XX	DNA encoding native BONT serotype A (BONTA) C-terminal fragment (Hc).		
XX	XX	Botulinum neurotoxin; heavy chain; BONT; serotype A;		
KW	KW	C-terminal fragment; Hc; Venezuelan equine encephalitis virus replicon;		
KM	KM	VEE; botulism; vaccine; diagnosis; drug screening; ds.		
XX	OS	Clostridium botulinum.		
XX	XX			
FH	FT	Location/Qualifiers		
FT	FT	1..1317		
FT	FT	/*tag= a		
FT	FT	/product= "Native botulinum neurotoxin serotype A		
FT	FT	(BONTA) heavy chain C-terminal fragment"		
FT	FT	/transl_except= (pos:1294..1314, aa:Gly)		
XX	PN	WO200002524-A2.		
XX	XX	20-JAN-2000.		
XX	XX			
PF	XX	09-JUL-1999; 99WO-US15570.		
XX	XX			
PR	XX	10-JUL-1998; 98US-0092416.		
PR	XX	12-MAY-1999; 99US-0133870.		
XX	XX			
PA	XX	(USNE-) US MEDICAL RES INST INFECTIOUS DISEASES.		
XX	XX			
PI	XX	Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;		
XX	XX			
DR	XX	WPI: 2000-160827/14.		
XX	XX	P-PSDB: AAY77142.		
PT	XX	Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum		
PT	XX	toxln serotypes A-G, is used for inducing an immune response against		
XX	XX	botulinum -		
PS	XX	Example 3; Page 52; 54pp; English.		
CC	XX	The invention relates to novel vaccines that induce a protective immune		
CC	XX	response against botulinum neurotoxin (BONT) serotypes A, B, C, D, E, F		
CC	XX	and G (BONTA-BONTG). The vaccine of the invention is novel recombinant		
CC	XX	DNA construct comprising a vector, and at least one nucleic acid		
CC	XX	fragment comprising a C-terminal heavy chain fragment (Hc) from BONT		
CC	XX	serotypes A-G. In preferred embodiments of the invention, the vector is a		
CC	XX	Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this		
CC	XX	vector results in the production of large amounts of a protein encoded by		

CC a sequence cloned into the replicon. The constructs are used to produce
 CC vaccines against botulism. The proteins can also be used as diagnostic
 CC tools for the diagnosis of botulism. The transformed host cells can be
 CC used to analyse the effectiveness of drugs and agents which inhibit toxin
 CC effects. The vaccine currently used against botulism is dangerous
 CC and expensive to produce, and contains formalin, which is very painful
 CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
 CC the 7 serotypes are represented in the formulation. The novel vaccine
 CC of overcomes these problems, as it is easily purified, and available in
 CC large quantities. It is also expressed in the lymph nodes for a better
 CC immune response. The present sequence represents DNA encoding native
 CC BoNTA heavy chain C-terminal fragment (Hc) used in an exemplification of
 CC the present invention.

CC
 XX
 SO Sequence 1317 BP; 560 A; 108 C; 211 G; 438 T; 0 other;

Query Match 54.8%; Score 726; DB 21; Length 1317;
 Best Local Similarity 72.4%; Pred. No. 2,4e-194;
 Matches 942; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

QY 19 tctacctcctgatacatcaagaacatcctacatcctcctggaacctcgcgtac 78
 DB 16 tctacattactgataatataagaatattatatactcttataattgaattagaat 75
 QY 79 gaatccaatcacctgacacctgtctgcgtacgcttccaacatccaatcgtttctaa 138
 DB 76 gaaagtatcatcttaataagacttctcgtatgcatccaataaataatattgtagtaa 135
 QY 139 gtaacttcgataccgatacgaagaatcgaatccagctgttcaatctgaaatctccaa 198
 DB 136 gtaaatcttgataccaatagataaataatccaattcattattagaagttagtaa 195
 QY 199 atcgaagtttcctcgaagaatgctatcgtatacaacctctatgtaacgaattctccac 258
 DB 196 atgagagtaatttaaaaaatgctatgataataatgtagtaagaatttagtaact 255
 QY 259 tccctcgtatccgtatccgaataactccaatccatctcctcgaacatgaatacacc 318
 DB 256 agcttcttgataagaatctcctaagatatttaacagtaagcttaataatgaatataca 315
 QY 319 atcatacactgcatgtaaaaaaatctcgttgytaagaatctcgtacactacgtgtaac 378
 DB 316 ataataattgtatgtaaaaaatcagatgaggaagtatcacttaattatgtaagaata 375
 QY 379 atcggactctgcagacactcagaagaatcaacagcgtgtgattcaataactctag 438
 DB 376 atccggactctacaggaactcaggaataaacaagaagtagtttaataacagtaaa 435
 QY 439 atgatacaactctgatacatcaatcgtcgtatcctcgttaccatcaccaacatcgt 498
 DB 436 atgtaataatatacagttataataacagatgatttttgaactatacctaataataga 495
 QY 499 ctgaataatccaataatcatatacaacgacctgtatgcagacgaacacgatactcaat 558
 DB 496 ttaataatccttaaatattatataatgtaagaattaaagtaacaaacaaatccaat 555
 QY 559 ctggtaacatccacgcttcttaataacatcgttcaacagcagcgtgtgcgtgacct 618
 DB 556 ttaagtataatctcagtgtaataataatgtaattgaattagatggtgtagagataca 615
 QY 619 caccgctacatctgatacaaatcatcaatcgttgcgaacaagactgaacaaaagaa 678
 DB 616 catagataatttgataaaaaatatttaactcttttgataagaattaaatgaaaaaagaa 675
 QY 679 atcaagaacctgtacgaacacagtcgaattctgtagtacctgaagaagcttctgggtgac 738
 DB 676 atcaagaattatataatgatacaatcaatcaggtattttaaaagacttctgggtgac 735
 QY 739 taactgagtagcagaacaacgcttactacatgctgaatcgtacgatacgaacaaatagct 798
 DB 736 tattacaataatgataacacatcactatagtaattatataatgatacgaataaataatgct 795

QY 799 gacgtcaaatgtagatccgcgttaccatgtaacctgaagaagtcgcgctgttctgt 858
 DB 796 gatgttaaatatgtagatagatagagattatgcttcttaagaaggcctcaggtagcgt 855
 QY 859 atgataccaacatctacgttaactcttcctcgtacggttgcacaaatcatcaag 918
 DB 856 atgataccaacatttatttaattcaagttgttagtaggggacaaaatttataaaa 915
 QY 919 aaatcgcgtctgtgaacaagaagacatatcgttcgcaacaaatgtagtataatcaat 978
 DB 916 aaatcgtcttgcgaataaagataatattgttagaataatgatacgtgtatataat 975
 QY 979 gttgtagtgaagaacaagaataccgtctgcgtaccacatgctctccagcgtgttagaa 1038
 DB 976 gtagtagttaaataaagaataatagtttagtactaataatgcacacagcgttagaa 1035
 QY 1039 aagatctgtcgtctcgtgaatacccgagcgttggtaactcgtctcagtagttagta 1098
 DB 1036 aaatcactaagtgacatgaagaataacatcgtatgtagaagaatcgaagtagtaag 1095
 QY 1059 aaatccaagaacgaccaggtatcactacaatgcaaatgaaatcgtcagacacaaat 1158
 DB 1096 aagtcataaaatgacaaaggaatacaataaagcaaatgaaatcttaagaataat 1155
 QY 1159 gtaacgatacgttctacatcgttccacaggtccaacatcgcctaaactgttgc 1218
 DB 1156 ggaagtatagataggttctatagatgattcatcagctttaaataatagcctaaactag 1215
 QY 1219 tccaactgtgtaacatcgtcagatcgaacgcttccctcgcgaactcgtgtccttggag 1278
 DB 1216 agtaattgtataataagacaaatagaagaatcctagtagaacttgggtcgtcagggaa 1275
 QY 1279 ttcatccggttgaagacgttgggtggaacgtccgcgtgta 1320
 DB 1276 ttatccctgtagatgagtgagggagaaagccacgtgtaa 1317

RESULT 14
 AA87218
 ID AA87218 standard; DNA; 2532 BP.
 AC AA87218;
 DT 08-MAY-2000 (first entry)
 XX
 DE DNA encoding native botulinum neurotoxin serotype A (BoNTA).
 XX
 KW Botulinum neurotoxin; heavy chain; BoNT; serotype A;
 KW Venezuelan equine encephalitis virus replicon;
 KW VEE; botulism; vaccine; diagnosis; drug screening; ds.
 XX
 OS Clostridium botulinum.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2532
 FT /tag= a
 FT /product= "Native botulinum neurotoxin serotype A
 FT (BoNTA) heavy chain"
 FT /transl_except= (pos:2509..2529, aa:GLY)
 FT
 PN WO200002524-A2.
 XX
 PD 20-JAN-2000.
 XX
 PE 09-JUL-1999; 99WO-US15570.
 XX
 PR 10-JUL-1998; 98US-0092416.
 PR 12-MAY-1999; 99US-0133870.
 XX
 PA (USME-) US MEDICAL RES INSTR INFECTIOUS DISEASES.
 XX
 PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
 XX

DR WPI: 2000-160827/14.
 DR P-PSDB: AAY77140.
 XX
 PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
 PT toxin serotypes A-G, is used for inducing an immune response against
 PT botulinum -
 XX
 PS Example 3: Page 49-50; 54pp: English.
 XX
 CC The invention relates to novel vaccines that induce a protective immune
 CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
 CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
 CC DNA construct comprising a vector, and at least one nucleic acid
 CC fragment comprising a C-terminal heavy chain fragment (HC) from BoNT
 CC serotypes A-G. In preferred embodiments of the invention, the vector is a
 CC Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
 CC vector results in the production of large amounts of a protein encoded by
 CC a sequence cloned into the replicon. The constructs are used to produce
 CC vaccines against botulinum. The proteins can also be used as diagnostic
 CC tools for the diagnosis of botulinum. The transformed host cells can be
 CC used to analyse the effectiveness of drugs and agents which inhibit toxin
 CC effects. The vaccine currently used against botulinum is dangerous
 CC and expensive to produce, and contains formalin, which is very painful
 CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
 CC the 7 serotypes are represented in the formulation. The novel vaccine
 CC of overcomes these problems, as it is easily purified, and available in
 CC large quantities. It is also expressed in the lymph nodes for a better
 CC immune response. The present sequence represents DNA encoding native
 CC BoNTA heavy chain used in an exemplification of the present invention.
 XX
 SO Sequence 2532 BP; 1038 A; 241 C; 404 G; 849 T; 0 other:
 Query Match 54.8%; Score 726; DB 21; Length 2532;
 Best Local Similarity 72.4%; Pred. NO. 3.3e-194;
 Matches 942; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

QY 559 ctgggtaacatccagctcttaatacatcagttcaaacctggagcgtgtgtgtacac 618
 DB 1771 ttggtaataatcattcgttgaataataataatgattcaaatgattgtgtgagataca 1830
 QY 619 caacgcttgcgtgataataactcaatcgtctgtgacaagaactggacaagaaga 678
 DB 1831 catagataatttggataaataatttcaatcttttggataaagaataataaagaaga 1890
 QY 679 atcaaaagcttgcagacaacagctccaaattcgtatccctgaagaacttcgtgggtgac 738
 DB 1891 atcaaaagattatagataataatcaatcaaaattcagattttaaagaactttgggtgac 1950
 QY 739 taactgcagtaagacaacacggtactacatgctggaactctgtagatccgacaatacgt 798
 DB 1951 taattcaatatgataaaccatactactatgtttaaatttatatgatacaataatgtgc 2010
 QY 799 gacgtcaaatgtagtatcccggttacaatgattaccctgaagaagtcgcgtgtgtgt 858
 DB 2011 gatgtataatgtagtattaggtattagattatgatacttaaaagcctagagtgatgta 2070
 QY 859 atgactacaacatactaccctgaactcttcctgttaccgtgtgataccaatcatalcaag 918
 DB 2071 atgactacaacacttatttatttaattcaagtttgtataggggacaattattataaaa 2130
 QY 919 aatatcgctgtgtgaacaagaacatatcgttgcacaacatgatacgtgtataacatcat 978
 DB 2131 aatatcgcttgcgaataaagaataatattgttgaataataatgatacgttgtatataat 2190
 QY 979 gttgtatgtaagaacaagaataacgctcgtgcgtacacatgcttcaagctgtgtgaga 1038
 DB 2191 gtatgattttaaataaagaataatagtttagctaacatgacacagcagcggtataga 2250
 QY 1039 aagatcttgcgtctgtggaatcccgagcgtgtgtaactctgtcaagttgtatgta 1098
 DB 2251 aaaaactaagtgatagtaataatactcgtgtgtaggaatactgaagtagtagtaacgt 2310
 QY 1099 aatccaagaacgacaggggtatcactacaacaatgcaaaatgaactgcagaacaacat 1158
 DB 2311 aagtcacaaaatgataaagaataaacaataaataatgcaaaatgaattacagaataat 2370
 QY 1159 ggtlaagatacgtgttcatcgtttccacagttcaacaatactgaactgaactgtgt 1218
 DB 2371 gggaaatgataagcctttagagattcactcagtttaataataatagctaaactagtag 2430
 QY 1219 tccaactgtgtacatcgttcagatcgaaacgttccctcgcacactcgtgtgtgtgtg 1278
 DB 2431 agtaattggtataatagacaataatagaaagatctagtaggaacttgggtcgtcatg 2490
 QY 1279 ttcactccggttgatgacggttgggtgtgaacgttcgtgtaa 1320
 DB 2491 ttatttccctgtatgattgattggtggaagaagccactgtaa 2532

RESULT 15
 AAT29244
 ID AAT29244 standard; DNA; 3891 BP.
 XX
 AC AAT29244;
 XX
 XX
 DT 09-JUL-1996 (first entry)
 XX
 DE C. botulinum type A neurotoxin gene.
 XX
 KW Toxin; neurotoxin; fusion protein; antitoxin; vaccine; immunogen; ds.
 XX
 OS Clostridium botulinum.
 XX
 PN W09612802-A1.
 XX
 PD 02-MAY-1996.
 XX
 PF 23-OCT-1995; 95WO-US13737.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:32:26 ; Search time 139.08 Seconds
(without alignments)
2341.890 Million cell updates/sec

Title: US-09-611-419a-5

Perfect score: 1326

Sequence: 1 gaattcgaacgatggcctc.....aacgtccgtctaagattc 1326

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1302.2	98.2	1330	1	US-08-480-604A-22	Sequence 22, Appl
2	1302.2	98.2	1330	2	US-08-405-496A-22	Sequence 22, Appl
3	1302.2	98.2	1330	4	US-08-915-136-22	Sequence 22, Appl
4	1302.2	98.2	1402	1	US-08-480-604A-25	Sequence 25, Appl
5	1302.2	98.2	1402	2	US-08-405-496A-25	Sequence 25, Appl
6	1302.2	98.2	1402	4	US-08-915-136-25	Sequence 25, Appl
7	726	54.8	3891	1	US-08-480-604A-27	Sequence 27, Appl
8	726	54.8	3891	2	US-08-405-496A-27	Sequence 27, Appl
9	726	54.8	3891	4	US-08-915-136-27	Sequence 27, Appl
10	166.4	12.5	1359	1	US-07-618-312A-3	Sequence 3, Appl
11	166.4	12.5	1359	1	US-08-280-228-3	Sequence 3, Appl
12	74.2	5.6	1359	1	US-07-618-312A-1	Sequence 1, Appl
13	74.2	5.6	1359	1	US-08-110-786A-7	Sequence 7, Appl
14	74.2	5.6	1359	1	US-08-280-228-1	Sequence 1, Appl
15	74.2	5.6	1359	1	US-08-668-381A-6	Sequence 6, Appl
16	60	4.5	5511	3	US-08-928-361B-2	Sequence 2, Appl
17	60	4.5	5511	3	US-08-928-361B-1	Sequence 1, Appl
18	57	4.3	7334	3	US-08-700-651-1	Sequence 1, Appl
19	57	4.3	5163	3	US-08-928-361B-4	Sequence 4, Appl
20	57	4.3	5318	3	US-08-700-651-2	Sequence 2, Appl
21	57	4.3	5318	3	US-08-928-361B-3	Sequence 3, Appl
22	40	3.0	1690	1	US-08-276-452A-24	Sequence 24, Appl
23	40	3.0	1690	2	US-08-798-744-24	Sequence 24, Appl
24	37.8	2.9	2277	1	US-08-676-967-2	Sequence 2, Appl
25	37.8	2.9	2277	1	US-08-676-967-2	Sequence 2, Appl
26	37.8	2.9	2277	2	US-09-098-487-2	Sequence 2, Appl
27	37.2	2.8	18596	4	US-09-318-448-11	Sequence 11, Appl

28	36.8	2.8	1430	1	US-08-276-452A-25	Sequence 25, Appl
29	36.8	2.8	1430	2	US-08-798-744-25	Sequence 25, Appl
30	34.6	2.6	2094	3	US-08-714-918-87	Sequence 87, Appl
31	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
32	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
33	34.6	2.6	2094	4	US-09-265-315-87	Sequence 87, Appl
34	34.4	2.6	951	1	US-07-855-412B-2	Sequence 2, Appl
35	34.4	2.6	951	2	US-08-308-887A-2	Sequence 2, Appl
36	34.4	2.6	951	3	US-08-881-094-2	Sequence 2, Appl
37	34.2	2.6	198	5	PCT-US95-10668-1	Sequence 1, Appl
38	34.2	2.6	198	5	PCT-US95-10668-2	Sequence 1, Appl
39	34.2	2.6	198	5	PCT-US95-10668-3	Sequence 1, Appl
40	34.2	2.6	198	5	PCT-US95-10668-4	Sequence 1, Appl
41	34.2	2.6	3292	1	US-07-814-964-12	Sequence 12, Appl
42	34.2	2.6	3292	1	US-08-258-442-12	Sequence 12, Appl
43	34.2	2.6	3292	1	US-08-328-809-7	Sequence 7, Appl
44	34.2	2.6	3292	5	PCT-US92-11107-12	Sequence 12, Appl
45	34.2	2.6	7218	1	US-08-232-463-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-480-604A-22
; Sequence 22, Application US/08480604A
; Patent No. 5736139
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHAYE, NISHA V.
; APPLICANT: FIRCA, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; PREVENTION OF C. DIFFICILE DISEASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,604A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422,711
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,496
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPHD-01763
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1350 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1314
 US-08-480-604A-22

Query Match 98.2%; Score 1302.2; DB 1; Length 1330;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

14 tggcctctactcttactgatacatcaagaacatcaatcaatccatccatccgaactcgc 73
 11 tgcgtcttactcttactgatacatcaagaacatcaatccatccatccgaactcgc 70
 74 gctcgaatccaatcaactgatacgcgttctcgtcgaactccaaatcaatcgcgtt 133
 71 gctacgaatccaatcaactgatacgcgttctcgtcgaactccaaatcaatcgcgtt 130
 134 ctaaatgtaactgatacgcgttctcgtcgaactccaaatcaatcgcgttctcgtt 193
 131 ctaaatgtaactgatacgcgttctcgtcgaactccaaatcaatcgcgttctcgtt 190
 194 ccaaaatcgaagttaactcgaagaatgtaactgatacgaactcgaagaactcct 253
 191 ccaaaatcgaagttaactcgaagaatgtaactgatacgaactcgaagaactcct 250
 254 caactcctcttctggaatcgatccgaagaatcaactcgaactcctcgaagaat 313
 251 caactcctcttctggaatcgatccgaagaatcaactcgaactcctcgaagaat 310
 314 acacacatcaactgatacgcgttctcgtcgaactcgaagaatcgatccgaagaat 373
 311 acacacatcaactgatacgcgttctcgtcgaactcgaagaatcgatccgaagaat 370
 374 aaatcactggaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 433
 371 aaatcactggaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 430
 434 ctcgaatgatacaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 493
 431 ctcgaatgatacaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 490
 494 atcgtctgatacaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 553
 491 atcgtctgatacaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 550
 554 ccaatctggaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 613
 551 ccaatctggaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 610
 614 acacacatcaactgatacgcgttctcgtcgaactcgaagaactcgaagaactcgttgaat 673
 611 acacacatcaactgatacgcgttctcgtcgaactcgaagaactcgaagaactcgttgaat 670
 674 aagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 733
 671 aagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 730
 734 gtaactcgtcgaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 793
 731 gtaactcgtcgaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 790

794 acgttgaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 853
 791 acgttgaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 850
 854 cgttgaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 913
 851 cgttgaactcgaagaactcgaagaatcaactcgaagaactcgttgaat 910
 914 tcaagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 973
 911 tcaagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 970
 974 tcaagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1033
 971 tcaagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1030
 1034 tgaagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1093
 1031 tgaagaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1090
 1094 taatgaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1153
 1091 taatgaatcaagaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1150
 1154 acaatgtaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1213
 1151 acaatgtaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1210
 1214 tgcctcgaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1273
 1211 tgcctcgaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1270
 1274 ggaatgtaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1320
 1271 ggaatgtaactcgtcgaagaactcgaagaatcaactcgaagaactcgttgaat 1317

RESULT 2
 US-08-405-496A-22
 Sequence 22, Application US/08405496A
 Patent No. 5919665
 GENERAL INFORMATION:
 APPLICANT: WILLIAMS, JAMES A.
 TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 MONTGOMERY STREET, SUITE 2200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/405,496A
 FILING DATE: 16-MAR-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791


```

1 GENERAL INFORMATION:
2 APPLICANT: WILLIAMS, JAMES A.
3 TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
4 TITLE OF INVENTION: NEUROTOXIN
5 NUMBER OF SEQUENCES: 30
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: MEDLEN & CARROLL, LLP
8 STREET: 220 MONTGOMERY STREET, SUITE 2200
9 CITY: SAN FRANCISCO
10 STATE: CALIFORNIA
11 COUNTRY: USA
12 ZIP: 94104
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patent Release #1.0, Version #1.30
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/405,496A
21 FILING DATE: 16-MAR-1995
22 CLASSIFICATION: 424
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 08/329,154
25 FILING DATE: 25-OCT-1994
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US 08/161,907
28 FILING DATE: 02-DEC-1993
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 07/985,321
31 FILING DATE: 04-DEC-1992
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: US 07/429,791
34 FILING DATE: 31-OCT-1989
35 ATTORNEY/AGENT INFORMATION:
36 NAME: INGOLIA, DIANE E.
37 REGISTRATION NUMBER: 40,027
38 REFERENCE/DOCKET NUMBER: OPHD-01308
39 TELECOMMUNICATION INFORMATION:
40 TELEPHONE: (415) 705-8410
41 TELEFAX: (415) 397-8338
42 INFORMATION FOR SEQ ID NO: 25:
43 SEQUENCE CHARACTERISTICS:
44 LENGTH: 1402 base pairs
45 TYPE: nucleic acid
46 STRANDEDNESS: double
47 TOPOLOGY: linear
48 MOLECULE TYPE: DNA (genomic)
49 FEATURE:
50 NAME/KEY: CDS
51 LOCATION: 1..1386
52
53 US-08-405-496A-25

```

Query Match	98.2%	Score 1302.2	DB 2	Length 1402
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1304	Conservative	0	Mismatches 3	Indels 0
Qy	14	tgagctctacctcactgatacatcaagaacatcatcaatccctccatccctgaacctgc	73	
Db	83	tgctgtctaaccttccactgatatcatcaagaacatcatcaatccctccatccctgaacctgc	142	
Qy	74	gctacgaatccaatcaactgatactgcacctgctctgcctgaagcttccaaatcaaacatcggtt	133	
Db	143	gctaacgaattccatcaccctgacgcacctgtccgtacacctccaaatcaaacatccggtt	202	
Qy	134	ctaaaggttaactcgatccgatccgcacagaatccagatccagcttgttccaatctggaatctt	193	
Db	203	ctaaaggttaactcttcgatccgatccgcacagaatccagatccagcttgttccaatctggaatctt	262	
Qy	194	ccaaaatcgaaattatccttgaaagatgcatctgatalacaactctatgtacgaanaacttct	253	
Db	263	ccaaaatcgaaattatccttgaaagatgcatctgattacaaactctatgtacgaanaacttct	322	

OY	254	ccacctcctcttgatccgfatcccgaaatcactcaacccactctctctcgaaacaatgat	313
Db	323	ccacctctcttctggatccgfatcccgaaatcactcaacccactctctctcgaaacaatgat	382
OY	314	aacaccatcatcaactgatctgaaacaactcttgcttggaaglatctctgaactcagctg	373
Db	383	acaccatcatcaactgatctgaaacaactcttgcttggaaglatctctgaactcagctg	442
OY	374	aaatcaatcttgacctcgcgaggacaatcgaagaaatcaacaacggtgttgtatccaatct	433
Db	443	aaatcaatcttgacctcgcgaggacaatcgaagaaatcaacaacggtgttgtatccaatct	502
OY	434	ctcagaatgaatcaactctcgaactcaaatcgcgcggaatcttgcttaccatccacaaca	493
Db	503	ctcagaatgaatgaatcaactctcgaactcaaatcgcgcggaatcttgcttaccatccacaaca	562
OY	494	atcgtctgaataactccaanaatcacaacgagccgcttgatcgaccagaacccgact	553
Db	563	atcgtctgaataactccaanaatcacaacgagccgcttgatcgaccagaacccgact	622
OY	554	ccaatctgggttaaacatccacgcttctaatacaatgttcaactcgaacggtgtgtgt	613
Db	623	ccaatctgggttaaacatccacgcttctaatacaatgttcaactcgaacggtgtgtgt	682
OY	614	acaactcacgcgtatcatcttgatcaaatcaactcaatctgtctcgacaagaacggaacga	673
Db	683	acaactcacgcgtatcatcttgatcaaatcaactcaatctgtctcgacaagaacggaacga	742
OY	674	aagaatccaagaacccgtgacgaacaaccgactccaatctcgtgatacctgaagaactctggg	733
Db	743	aagaatccaagaacccgtgacgaacaaccgactccaatctcgtgatacctgaagaactctggg	802
OY	734	gtgcactacctcgaatgaacaacaaccgactcaaatctgtgaatctgtaacgataccgaacaat	793
Db	803	gtgcactacctcgaatgaacaacaaccgactcaaatctgtgaatctgtaacgataccgaacaat	862
OY	794	acgttgacgctcaacaacgtgtagtataccgcggtgtacatgtaactcgtgaacgaagctccgcgtg	853
Db	863	acgttgacgctcaacaacgtgtagtataccgcggtgtacatgtaactcgtgaacgaagctccgcgtg	922
OY	854	ctgtgatagaactcaacaatacttaccctgaactctcccttgtaacggtgtgataccaatcatca	913
Db	923	ctgtgatagaactcaacaatacttaccctgaactctcccttgtaacggtgtgataccaatcatca	982
OY	914	tcagaanaatacgcgtctcgtgtgaacaagaagacatactgtcgcacaacaatgtaactcgtgtataca	973
Db	983	tcagaanaatacgcgtctcgtgtgaacaagaagacatactgtcgcacaacaatgtaactcgtgtataca	1042
OY	974	tcaatgtctgtagtttaagaacaagaataaccgcttggtctacccaatgcttctcaacgctggtg	1033
Db	1043	tcaatgtctgtagtttaagaacaagaataaccgcttggtctacccaatgcttctcaacgctggtg	1102
OY	1034	tagaanaagatctgtctgcctctggaatatcccggaactgtgtaactcgtctcgaagtagtg	1093
Db	1103	tagaanaagatctgtgtctgcctctggaatatcccggaactgtgtaactcgtctcgaagtagtg	1162
OY	1094	taatgaanaatccaagaacgaccgaggtatacaatacaaatgtcaanaatgtaactcgaagaca	1153
Db	1163	taatgaanaatccaagaacgaccgaggtatacaatacaaatgtaactcgaagaca	1222
OY	1154	acaatgtgaagaagatactggttatacgtgttccaacaagttaaanaatatcgtataactcgtg	1213
Db	1223	acaatgtgaagaagatactggttatacgtgttccaacaagttaaanaatatcgtataactcgtg	1282
OY	1214	ttgcttccaactgtgtaacaatcgttcgaatcgaaacggttccctctcgacaactcgtggttgtctt	1273
Db	1283	ttgcttccaactgtgtaacaatcgttcgaatcgaaacggttccctctcgacaactcgtggttgtctt	1342
OY	1274	ggagagttcaatcccggttgatgaacggtgtgggtgtgaacgctcgcgtgtgaa	1320
Db	1343	ggagagttcaatcccggttgatgaacggtgtgggtgtgaacgctcgcgtgtgaa	1389


```

1      RESULT      6
2      US-08-915-136-25
3      ; Sequence 25, Application US/08915136
4      ; Patent No. 6290360
5      ; GENERAL INFORMATION:
6      ; APPLICANT: KINK, JOHN A.
7      ; APPLICANT: THALLEY, BRUCE S.
8      ; APPLICANT: PADHTE, NISHA V.
9      ; APPLICANT: FIRCA, JOSEPH R.
10     ; APPLICANT: STAFFORD, DOUGLAS C.
11     ; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
12     ; TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
13     ; NUMBER OF SEQUENCES: 32
14     ; CORRESPONDENCE ADDRESS:
15     ; ADDRESSEE: MEDLEN & CARROLL, LLP
16     ; STREET: 220 MONTGOMERY STREET, SUITE 2200
17     ; CITY: SAN FRANCISCO
18     ; STATE: CALIFORNIA
19     ; COUNTRY: UNITED STATES OF AMERICA
20     ; ZIP: 94104
21     ; COMPUTER READABLE FORM:
22     ; MEDIUM TYPE: Floppy disk
23     ; COMPUTER: IBM PC compatible
24     ; OPERATING SYSTEM: PC-DOS/MS-DOS
25     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
26     ; CURRENT APPLICATION DATA:
27     ; APPLICATION NUMBER: US/08/915,136
28     ; FILING DATE:
29     ; CLASSIFICATION:
30     ; PRIOR APPLICATION DATA:
31     ; APPLICATION NUMBER: 08/480,604
32     ; FILING DATE:
33     ; PRIOR APPLICATION DATA:
34     ; APPLICATION NUMBER: US 08/405,496
35     ; FILING DATE: 16-MAR-1995
36     ; PRIOR APPLICATION DATA:
37     ; APPLICATION NUMBER: US 08/329,154
38     ; FILING DATE: 25-OCT-1994
39     ; PRIOR APPLICATION DATA:
40     ; APPLICATION NUMBER: US 08/161,907
41     ; FILING DATE: 02-DEC-1993
42     ; PRIOR APPLICATION DATA:
43     ; APPLICATION NUMBER: US 07/985,321
44     ; FILING DATE: 04-DEC-1992
45     ; PRIOR APPLICATION DATA:
46     ; APPLICATION NUMBER: US 07/429,791
47     ; FILING DATE: 31-OCT-1989
48     ; ATTORNEY/AGENT INFORMATION:
49     ; NAME: INGOLIA, DIANE E.
50     ; REGISTRATION NUMBER: 40,027
51     ; REFERENCE/DOCKET NUMBER: OPHD-01763
52     ; TELECOMMUNICATION INFORMATION:
53     ; TELEPHONE: (415) 705-8410
54     ; TELEFAX: (415) 397-8338
55     ; INFORMATION FOR SEQ ID NO: 25:
56     ; SEQUENCE CHARACTERISTICS:
57     ; LENGTH: 1402 base pairs
58     ; TYPE: nucleic acid
59     ; STRANDEDNESS: double
60     ; TOPOLOGY: linear
61     ; MOLECULE TYPE: DNA (genomic)
62     ; FEATURE:
63     ; NAME/KEY: CDS
64     ; LOCATION: 1..1386
65     ; US-08-915-136-25

```

Query Match	98.2%	Score 1302.2	DB 4	Length 1402
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1304	Conservative	0	Mismatches 3	Indels 0
Gaps				0

0Y 14 tggccttacccttcactgtaataatcaagaacatcataataactccatcctgaactgc 73
 |||||||

D	83	TGCTGTACTCTTACGTACGTAATACATCAAGAACATCATCAATACCTGCATCCGTGAACCTGGC	142
Q	74	gctagaatcccaatccaccgtatcgaccgtgtctcgtaacgtccccaatacaacatcggtt	133
D	143	GCTAGAAATCCAAATCACCCTGATGACCTGCTGCTACGCTTCCAAAATCAACATCGGGT	202
Q	134	ctaaagttaacttcgatccgatcgacaagaatcaatgatcagctgttcaatctgaaactt	193
D	203	CTAAAGTTAACTTGTGATCCGATCCGACACAAAGATAGATTCAGCGTTCATTCGGAATCTT	262
Q	194	ccaatatgaagttatccttgaagaatgctctgtatcaaaccttaagtacgaagaactct	253
D	263	CCAAATATGAAGTTATCTCTGAGAGATGCTATGCTATACACTCTATATACGAAACTTCT	322
Q	254	ccaactcccttcgtgatccgtatcccgaaatcaacttcaactcctctctcgtgaacaatgat	313
D	323	CCACCTCCTTCTGGATCCGATATCCGAATATCTTCAATTCATCTCTGAAACATGAAAT	382
Q	314	acaccatcaatcaactcgatggaagaaacaaattctggtttggaagtatctctgaactacggtg	373
D	383	ACACATCATATCACTGGATGAGAAACAAATTCCTGGTTGGAAGATATCTCGAACTACGGTG	442
Q	374	aaatcaatctggaactctgcagaaactcaagaagaatcaaacagcgttgtttatcaaatct	433
D	443	AAATCATTTGGACTCTCGAGGACCTTAGAATACTAAACAGCGTGTGTTATTCAAATACT	502
Q	434	ctcagatgatcaaatctcttgaactcacaatcaatctgcgtgtgatacttcgttaccatcacaca	493
D	503	CTCAATGATATCAATCTGTGACTACATCAATCGTGAGCTTGATCTCGTTACCATCCACACA	562
Q	494	atcgctcgaataactccaataatctacatcaacggcgctgtatcgaccgaagaacgatct	553
D	563	ATCGCTGGAATTAACCTCCAAAATCTACATCAACCGCGCTGTGATCGACCGAAACCGACT	622
Q	554	ccaactctgggttaaatccacagcttctcaataacatgatttcaaacctgacggctgttggt	613
D	623	CCAATCTGGGTAACATCCACGCTTCTATATACATCATGTTCCAACTGGACGGTTGTCTGTG	682
Q	614	acaactcacgcgtataatcttgatcaaatacttcaatctgttctgcacaagaactgaaacgaa	673
D	683	ACACTCACCGCTACATCTGGATCGTAATTAATTAATCTGTTCCGACAAAGAACTGAACGAA	742
Q	674	aagaaatcaaaagacctgaacgaacacagttccaaattctggtatcctgaaagactctggg	733
D	743	AAGAAATCAAAAGACTCTGACGACACACGATCCAAATCTGGTATCTGGAAGAACTTCTGGG	802
Q	734	gtgactaccctcagtaagaacaaacgfractacagtgctgtaattctgtacagatccgaacaaat	793
D	803	GTGACTTACTCTCACTAGACAAACCGTACTACATGCTGAATCTGTACGATCCGAAACAAAT	862
Q	794	acgttgacgtcaaaatgtagtgatacccgcggttatactacactgaaaggttccgcgttggt	853
D	863	ACGTTGAGCTCAACAATGTAGTATCCGGGGTTACATGTATACGTGAAGATGTCGGCGGT	922
Q	854	ctgttatagaactaccaacatctactgaactcttccctgtacccgttggttaccaaattcatca	913
D	923	CTGTATACTACCAACATCATCTACTCGAACTTCCCTGTACCGTGGTACTCAAAATTCATCA	982
Q	914	toaagaatatcgctcgtgtgaataagaacaaatctgttcgaacaaatggtgtgtatata	973
D	983	TCAAAATAATACGCGTCTGGTGAACAGGACAAATATCGTTGGCAACAATGTATGTTATCA	1042
Q	974	tcaatgttgtagtgtaagaacaaagaatcacgcgtctggtctacaagtctctcagagctgggt	1033
D	1043	TCAATGTTGTAGTTAAGAACAAAGAAATACCGCTCTGGCTACCAATGCTTTCAGGGCTGGTG	1102
Q	1034	tagaagaatctctgtcgtctggaatacccggaacgltggttaatctgttctcagtagtctg	1093
D	1103	TAGAAAGATCTTGTCTGCTGTGAAATCCCGGACGTGGTATCTGCTCAGTAGATTG	1166
Q	1094	taatgaaatccaagaacgcaccaggtgatacctaacaatgacaaatgaaatctgcagga	1153
D	1163	TAAATGAATTCACAAACGACGAGGTATCTCTAAACAAATGAATCTGCGAGGACA	1222

Db 3490 AAATATGCTTCTCGAAATAAGATATATTTGTAAGAAATAATGATCGTATATATTAAT 3549
Qy 979 gtluaftlaagaagaataacacgctctggtctaccaatgtctccagcgctggttagaa 1038
Db 3550 GTAGAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3609
Qy 1039 aagatctgtctgtctggaatacccgagcgttggtaactctgtccagagtaactgaa 1098
Db 3610 AAAATACTAAGTGCATGTGAATAACCTGATGTAGAAATCTAAGTCAAGTACTACTAATG 3669
Qy 1099 aaatccaaagacgacgaggtatctactaaatgcaaatctgcaaggaacaacat 1158
Db 3670 AAGTCAAAAATGATCAAGGAATAACAATTAATCAAAATGCAAAATTTCAAGATTAATAT 3729
Qy 1159 ggtacagatcagctgttccatcaggtctccacagttcaacaatatgtcctaaactgttct 1218
Db 3730 GGGATGATATAGCGTTTATAGGATTCATCGTTTAAATATAGCTAAACTAGTACGA 3789
Qy 1219 tccacactgtacaaatcgtaacatcgatccctctcgcacactgtggtctgtctgagag 1278
Db 3790 AGTATGTTGTTAATATAGCAAAATAGAAAGATCTAGTAGGACTTTGGTTGCTCATGGGA 3849
Qy 1279 ttcatcccggttgtagacggttgagggtgaacgctccgtgtaa 1320
Db 3850 TTTATCTCTGTAGATGATGATGGGAGAAAGCCACTGTAA 3891

RESULT 8

US-08-405-496a-27
Sequence 27, Application US/08405496a
Patent No. 5919665

GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.

TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP

STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO

STATE: CALIFORNIA
COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/405,496a
FILING DATE: 16-MAR-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154

FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.

REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01308

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 3891 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3888
US-08-405-496a-27

Query Match 54.8%; Score 726; DB 2; Length 3891;
Best Local Similarity 72.4%; Pred. No. 1.3e-197;
Matches 942; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

Qy 19 tctaccttaactgtaatacatcaagaacatcatcaactcctcaactcctgaactcgtctac 78
Db 2590 TCTACATTTACTGGAATATATTAAGAAATATTAATTAATTAATTAATTAATTAAT 2649
Qy 79 gaatccaatcaccttgtagcactgtctgcgtcagcttccaaataacacgtctctaa 138
Db 2650 GAAAGTAAATCAATTAATTAAGTATATCTAGTATGATCAAAATAATTAATTTGTA 2709
Qy 139 gttaaacttcgactcgatcgacagaaatcgatccagctgtcgaatcgtggaatctccaa 198
Db 2710 GTAAATTTTGATGCATATATTAATAAATCAATTTCAATTTAATTTAAGAAGTATAA 2769
Qy 199 atcgaagtataccctgaagaatgctatcgatcaactatgtagaagaactctccacc 258
Db 2770 ATTGAGGTAAATTTTAAAAATCTATTGTATTAATAGTATGATGAATAATTTTACT 2829
Qy 259 tcccttgtagatcgatcccgaaatactcaactccatctctcgtgaagaatgaatcacc 318
Db 2830 ACCTTTTGGATTAAGAAATTCCTAAGTTTAAACAGTAACTAATTAATGAATATTA 2889
Qy 319 atcatcaactcgatgaaacaactctcgttggaagatctctgtaactcgtggaatc 378
Db 2890 AATAATTAATTTGATGAATAAATATTTAGAGATGGAAGATGCACTTAATTTATGTA 2949
Qy 379 atctggaactcgcaggaacatcaagaatcaaacagcgtgtgtatctcaatactccag 438
Db 2950 ATCTGGACTTTACAGGATCTCAGGAAATAAACAAGAGATGTTTAAATACAGTCAA 3009
Qy 439 atgatacaatctctgactcaactcaactcgtgactcttgatccatcccaactcgt 498
Db 3010 ATGATTAATATATATCAATTAATTAACAGATGAGATTTTGTAACTTAATTAATA 3069
Qy 499 ctgaataactccaaatctcaactcaacgagcgtctgtagcgaagaacccgactcccat 558
Db 3070 TTAATTAATCTTAATAATTTATTAATTAATGAAGATTAATGATCAAAACCAATTT 3129
Qy 559 ctggtgtaacatccagctcttaataacatcaatcaatcaatcaatcaatcaatcaat 618
Db 3130 TTAGGTAATATTCATGCTAGTATATTAATTAATTAATTAATTAATTAATTAATTA 3189
Qy 619 caccgctaatcttgatcaactcaactcaactcgttctgcaagaagaactgcaagaaga 678
Db 3190 CATAGATATATTTGATTAATAATTTTAACTTTTGAATGAAGAAATTAATTAATA 3249
Qy 679 atcaagaactcgttagacaacccagctcgaatcttgatctcgtgaagaactcgtggtgac 738
Db 3250 ATCAAGAATTTATGATATATCAATCAATTTAAGGATTTTAAAGACTTTTGGGGTAT 3309
Qy 739 tacctgtagtgcgaacacgctactaactcgtcgaactcgttagcgtcgaacacacgt 798
Db 3310 TATTTAATATATGATTAACCATATATATGTTAAATTTATGATCAATTAATTAATG 3369
Qy 799 gacgtcaaaatgtagtgcgtccggttagatcgttagcgtgaagaagtcggtgtctgt 858
Db 3370 GATGTAATTAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3429
Qy 859 atgactaccaaatcactcgtgaactctccctgtagcgtgtagcgaatcactcatcaag 918


```

Db 589 GGTCTGGGCGCTATCCGTGAGCAACACATCTTAAAGCTGACCGTTGCACAC 648
QY 616 actaccgctacatactggaataactcaatctgtcgacaagaactgaacga 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 AACACCAGTACGTATCCAGACAGAGTTCGGTATCTTGCAAGACGACCCGAA 708
QY 676 gaatacaagaactgtgacgaacccggtccaaatctggttccctgaagaactc 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 GAGATGAAAAAGTGTATACACACTCTCTATACCTCTGCTGCTGCTGCGGT 768
QY 736 gactacctgagtagcaaacaccgtactacatgtgaatctgtacgacgcgaac 795
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 769 AACCCGCTGGCTTACGACACCGCAATATTAACCTGATCCCGTACTCTTAAGAC 828
QY 796 gttgacgtacaacatgtaaglatccgcygttacatgtacgtgaaggtccggt 855
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 829 GTTACACTGAAAAA-----CATCAGTACTGACTGACTGACCAACGGCGCTCTAC 882
QY 856 gttatgactaacacatctccgaactcttccctgtacgtgtgtacaaatcattc 915
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 ACTAACGTTAACTGAACATCTACTACCGACGCTGTGTACAAAGCGCTGAATTCTATC 942
QY 916 aagaatacgcgtctgtgaacaagacaat 945
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 943 AAACGCTACACTCCGAACACGAATCGAT 972

```

RESULT 11

US-08-280-228-3

Sequence 3, Application US/08280228

Patent No. 5571694

GENERAL INFORMATION:

APPLICANT: Makoff Dr, Andrew J
 APPLICANT: Romanos Dr, Michael A
 APPLICANT: Clare Dr, Jeffrey J
 APPLICANT: Fairweather Dr, Neil F
 TITLE OF INVENTION: VACCINES
 NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHAYE P.C.

STREET: 1100 No. 5571694th Glebe Road

CITY: Arlington,

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,228

CLASSIFICATION: 435

FILING DATE: 25-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/618,312

CLASSIFICATION: 435

FILING DATE: 27-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 8926832.0

FILING DATE: 28-NOV-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9006097.1

FILING DATE: 17-MAR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Wilson, Mary J.

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 117-163

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

TELEX: 200797 NIXN UR

INFORMATION FOR SEQ ID NO: 3:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Clostridium tetani
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1356
; US-08-280-228-3

Query Match      12.5%; Score 166.4; DB 1; Length 1359;
Best Local Similarity 52.3%; Pred. No. 5.2e-38;
Matches 486; Conservative 0; Mismatches 411; Indels 33; Gaps 4;

QY 43 aacatcatalaactccatccatcgtgaacgtcgtgacgaatccatccatcgtgacctg 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49 ATCCTGAAAAAGTCTACCATTTGAACTTGACATCAACACGATATTATCTCCGACATC 108
QY 103 tctcgctacgcttccaaatacaacatcggttccaaagttaacttgatccg---atgcac 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TCTGGTTTCAACTCCTCTGTTATCATATTCAGATCTCAATTGGTGGCGGCATCAAC 168
QY 160 aagaatcagatccagcgttccatcgtgaatcttccaaatcgaaagtatccatgaagaat 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 GGCNAAGCTATCCACCTCGGTTAACACGAATCTTGTGAATTCGTGCAACAGGCCATG 228
QY 220 gctatcgtatcaactctatgtagaanaacttccacccctcctcgtgacgtatccg 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 GACATCGAATACAAAGCAGCATGTTCAACAACTTCCGTTAGTCTTGGCTGCGCGTTCCG 288
QY 280 aaatacttaactccatc-----tctcgtgaacatgaataacacatcattcaactgc 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 AAAAGTTTCTGCTTCCACCTGGAACAGTACGGGACTAAGCAGTATCCATCATCAGTCT 348
QY 331 atggaanaacat-----tctggttggaagtatctctgaactcgtgtgaa 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 ATGAGAAACACTCCCTGCTCATCGGCTCTGTTGGTCTTTCCTGGAAGGTTAAACAC 408
QY 376 atcatcgtgaactctgcaagcaactcaggaatcaaacagcgtgtgtatcaataactct 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 CTGATCTGACCTGGAAGACTCCGGGGGAGATTGCTGAGATCACTTCCGAGACTG 468
QY 436 caagatgatacaatctctgactacatcaatcgtcgtgattctgttaccatccacaacat 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 CCGGACAAAGTTCAACGCGTACCTGGCTTAAACAAATGGGTTTTCATCACTATCACTA 528
QY 496 cgtcgtgaataactccaaatctacatacagcgctgtgacgaacgaacgatctcc 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 GGTCTGTCTTCTCTAACCCTGTATCAACAGCGGCTTGTGATGGCTCCGCTGAATACAT 588
QY 556 aatcgtgtaacatctcagcgttcttaataacatglttcaaatgacggtgtcgtgac 615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 589 GGTCTGGGCGCTATCCGCTGAGGACAAACATCATCACTTAAAGCTGACCGTTCACACAC 648
QY 616 actcaccgctacatactggaataactcaatctgttgcagaagaactgaacgaagaa 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 AACACCAGTACGTATCCAGACAGAGTTCGGTATCTTGCAAGACGACCCGAA 708
QY 676 gaatacaagaactgtgacgaacccggtccaaatctggttccctgaagaactc 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 GAGATGAAAAAGTGTATACACACTCTCTATACCTCTGCTGCTGCTGCGGT 768
QY 736 gactacctgagtagcaaacaccgtactacatgtgaatctgtacgacgcgaac 795
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 769 AACCCGCTGGCTTACGACACCGCAATATTAACCTGATCCCGTACTCTTAAGAC 828
QY 796 gttgacgtacaacatgtaaglatccgcygttacatgtacgtgaaggtccggtgttct 855

```

Db 829 GTTCACCTGAAAA-----CATCACTGACTACATGTACTGACCAACCGCGCTCTAC 882
Qy 856 gtatagaccaccaacacttaactgaactctccctgtaaccgtygtacaaatcatc 915
Db 883 ACTAAGCGTAACTGAACTCTACTACCGACGCTCTGTACAAAGCGCTGAAATTCATCATC 942
Qy 916 aagaatacgcgtctgtaaccagaacat 945
Db 943 AACGCTACACTCCGAAACACGAATGAT 972

RESULT 12
US-07-618-312A-1

; Sequence 1, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makolf Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1359 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Clostridium tetani
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1356
; US-07-618-312A-1

Query Match 5.6%; Score 74.2; DB 1; Length 1359;
Best Local Similarity 47.3%; Pred. No. 1.1e-11;
Matches 223; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

Qy 343 tctggttgaaagatctctgaactcagcgtgaaatcatctggaactcgcaggaactcag 402
Db 376 TCTGTTGGAGTCTATCACTTAAGTAACTTAATGACTTTAAAGATTCCGGG 435
Qy 403 gaaatcaaacagcgtgltglatcaaatcactcagatgatcaacaactctgactac 462
Db 436 GGAAGATTAGACAAATACTTTAGGATTTACTGATTAATTTATCTCTATTATGCA 495
Qy 463 aatcgctgtaactctggtccaccacacaacatcgctgaataaccacaaatctcac 522
Db 496 AATTAATGGGTTTATTAATCTATTACTAATGATGATTAATCTCTCTAATTTGATTA 555
Qy 523 aacgcgcctcgtacgacgaacacgcctccaatcgtgtaacatccacgcctta 582
Db 556 AATGAGACTACTTATGGGAAGTGCAGAAATTACTGCTTTAGACTTTAGAGAGTAA 615
Qy 583 aacatcgttcaactggaactgagcgtgltgltgtaactcaccgtacatctgataca 642
Db 616 AATATTAACATTTAAACTAGATGATTAATTAATTAATCAATACCTTTCTATTGATTA 675
Qy 643 tcaactcgttcgaagaagactggaagaaagaatcaagaactgtaagcaacccg 702
Db 676 TTTAGATATTTTGCAGAGCATTAATCCAAAGAGATGAAATTAATACACACTTA 735
Qy 703 tccaattcgtataccgtgaagactctggtggtgtaactcgtcagtaagcaacgtac 762
Db 736 TTAATCAATACCTTTTAAGAGACTCTGGGGAACCTTTAGATATGATACAGATTA 795
Qy 763 tacatcgtgaatctgtaagatccgaacaaatcagttgacgtcaacatgta 813
Db 796 TATTAATACAGACTACTCTACTTAAGATGTTCAATTAATAATAATA 846

RESULT 13
US-08-110-786A-7

; Sequence 7, Application US/08110786A
; Patent No. 5443966
; GENERAL INFORMATION:
; APPLICANT: FAIRWEATHER, Neil Fraser
; APPLICANT: MAKOLF, Andrew Joseph
; TITLE OF INVENTION: Expression of tetanus toxin fragment C
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 5443966th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,786A
; FILING DATE: 23-AUG-1993 1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,337
; FILING DATE: 29-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00943
; FILING DATE: 20-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8914122.0
; FILING DATE: 20 June 1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary J. Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000

QY 763 tacatgcgaatctgtacgacgcgaacaatacgtgtgacgcaacaatgta 813
|||
Db 796 TATTTAAATACAGTAGCTTCTAGTCTTAAGATGTTCAATGAAATAATA 846

RESULT 15
US-08-668-381A-6

: Sequence 6, Application US/08668381A
: Patent No. 5780024

GENERAL INFORMATION:

: APPLICANT: Brown, Robert H.
: APPLICANT: Fishman, Paul S.
: APPLICANT: Francis, Jonathan W.

: APPLICANT: Hosler, Betsy A.

: TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
: TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN

: NUMBER OF SEQUENCES: 6

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Fish & Richardson P.C.

: STREET: 225 Franklin Street

: CITY: Boston

: STATE: MA

: COUNTRY: USA

: ZIP: 02110-2804

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: OPERATING SYSTEM: IBM PC compatible

: SOFTWARE: Patentin Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/668,381A

: FILING DATE: 21-JUN-1996

: CLASSIFICATION: 514

: PRIORITY APPLICATION DATA:

: APPLICATION NUMBER: 60/000,473

: FILING DATE: 23-JUN-1995

: ATTORNEY/AGENT INFORMATION:

: NAME: Clark, Paul T.

: REGISTRATION NUMBER: 30,164

: REFERENCE/DOCKET NUMBER: 00786/269001

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 617/542-5070

: TELEFAX: 617/542-8906

: TELEX: 200154

: INFORMATION FOR SEQ ID NO: 6:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 1858 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: single

: TOPOLOGY: linear

: MOLECULE TYPE: DNA

: US-08-668-381A-6

Query Match 5.6%; Score 74.2; DB 1; Length 1858;

Best Local Similarity 47.3%; Pred. No.: 1.3e-11;

Matches 223; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 343 tctgttggaagatctgaatcgaatcgtgaaatcattcgtgacgacactcag 402
|||||
Db 875 TCTGTTGAGTGTACTTAAGTAACTTAATATGACCTTTAAAGATCCGGG 934
QY 403 gaatcaaacagcgtgtglatcaaatcctcaatgatacaatcctcgaactac 462
|||
Db 935 GGAGAGTTAGACAATTAATCTTTAGGATTTACCTGATTAATTAATCTTTAGCA 994
QY 463 aatgcgtgactctcgttacatccacaaatcgtcgaataactccaatctacatc 522
|||
Db 995 AATAAATGGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1054
QY 523 aacggcgtcgtgacgacgaaacgatactcgaatctcgtggaacacacgctccta 582
|||

Db 1055 AATGAGTACTTATGGAAGTGCAGAAATTAATGCTTTAGAGACTATTAGAGGATAT 1114
QY 583 aacatcatgtcaaacctgacggtgtcgtgacactcaacgctacatctgatacaatc 642
|||
Db 1115 AATTAATACATTAATAAATGATGATGATGATGATGATGATGATGATGATGAT 1174
QY 643 tcaatctgtcgaacgaactgacgaaagaagaatcaagaactgacgacacacag 702
|||
Db 1175 TTTAGATATTTTGCAGAAAGCATTAATCAAAAGAGATTGAAATTTATACAGATTAT 1234
QY 703 tccaattctglatccctgaagaactctggtgtgactcctcgtcagtaacgacacg 762
|||
Db 1235 TTATCTAATAACCTTTTAAAGAGACTCTGGGAAACCTTTACGATATGATACAGAAAT 1294
QY 763 tacatgcgaatctgtacgacgcgaacaatacgtgtgacgcaacaatgta 813
|||
Db 1295 TATTTAAATACAGTAGCTTCTAGTCTTAAGATGTTCAATGAAATAATA 1345

Search completed: September 16, 2002, 21:32:45
Job time: 15697 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 19:28:15 : Search time 4365.92 Seconds
(without alignments)
4099.242 Million cell updates/sec

Title: US-09-611-419A-5

Perfect score: 1326

Sequence: 1 gaattcgaaacgatagcctc.....aacgtccgctgaagattc 1326

Scoring table:

IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estrlo:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58.4	4.4	608	10	BI887904
2	56.2	4.2	1101	12	CNS00LOO
3	53.6	4.0	537	10	BM277996
4	53.6	4.0	551	10	BM278689
5	53.2	4.0	394	9	AU060224
6	52.6	4.0	447	10	BI594803
7	52	3.9	597	10	BM278181
8	51.8	3.9	500	10	BM278805
9	50.6	3.8	511	10	BM278805
10	50.4	3.8	569	10	BM278657
11	50	3.8	641	10	BM181884
12	49.6	3.7	645	9	AI389106
13	49.2	3.7	827	12	CNS02156
14	49	3.7	481	10	BR050030
15	48.8	3.7	628	9	AU060230
16	48.2	3.6	546	10	C22974
17	48.2	3.6	1101	12	CNS0100X

18	48	3.6	546	10	BF169335
19	48	3.6	556	10	BM277853
20	47.4	3.6	377	10	BI594814
21	47.4	3.6	450	9	AU060996
22	47	3.5	458	10	BI783045
23	47	3.5	518	10	BI782765
24	46.8	3.5	1175	10	BI872945
25	46.4	3.5	469	10	BF050033
26	45.4	3.4	386	10	BF050064
27	45.4	3.4	499	10	BM278692
28	45.4	3.4	536	10	BF050073
29	45.4	3.4	542	10	BM278786
30	45.4	3.4	554	10	C23753
31	45.2	3.4	501	10	BM277916
32	45.2	3.4	525	12	BM184460
33	45.2	3.4	525	12	CNS07PEA
34	45.2	3.4	1101	12	CNS000D1
35	45	3.4	564	10	BE606733
36	44.8	3.4	534	10	BM278087
37	44.6	3.4	1101	12	CNS016JY
38	44.4	3.3	522	10	BM278558
39	44.4	3.3	681	12	CNS02BOD
40	44.2	3.3	325	12	CNS03PK8
41	44.2	3.3	429	9	AU052930
42	44.2	3.3	500	12	B67199
43	44.2	3.3	519	10	BM278164
44	43.8	3.3	535	10	BI594948
45	43.6	3.3	443	10	BM039802

ALIGNMENTS

RESULT 1
BI887904
LOCUS
DEFINITION
Zf637-1-002159 zebrafish shield stage whole embryo cDNA library
MPMGp637 Danio rerio cDNA clone MPMGp637_10F4:MPMGp637F0410 5',
mRNA sequence.
ACCESSION
BI887904
VERSION
BI887904.1
KEYWORDS
zebrafish.
SOURCE
Danio rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 608)
Clark, M., Aanstad, P., Hennig, S., Johnson, S.L. and Lehrach, H.
EST sequencing of a zebrafish shield stage cDNA library normalised
by oligonucleotide fingerprinting
JOURNAL
Unpublished (2001)
COMMENT
Contact: Hennig S
Laboratory 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inhest. 63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
5' EST sequencing of clones from a zebrafish shield stage library,
normalised from 55,000 starting clones by oligonucleotide
fingerprinting
High quality sequence stop: 608.
FEATURES
source
1. 608
Location/Qualifiers
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="MPMGp637_10F4:MPMGp637F0410"
/clone_lib="zebrafish shield stage whole embryo cDNA
library MPMGp637"
/tissue_type="whole embryo"
/dev_stage="shield stage, 6 hrs post-fertilisation"
/lab_host="E.coli, XLI blue MRF"

BASE COUNT	207	a	218	c	41	g	140	t	2	others
ORIGIN	/note="vector: pSport1; Site_1: NotI; Site_2: SalI; oligo-dt-NotI primed, SalI adaptors, directionally cloned library normalised by oligonucleotide fingerprinting"									

Query Match	4.48;	Score 58.4;	DB 10;	Length 608;
Best Local Similarity	49.18;	Pred. No. 0.00035;		
Matches 181; Conservative	0;	Mismatches 187;	Indels 1;	Gaps 1.

QY	442	atcaacatctcgtacatatacaaa-cgcgtggatcttcgttaccatcaccacaatcgtct	500
Db	5	ATCAACTTCTCCAGGTACATCAACTCINTGTCCAGGTACATCAACTTCTCCAAACAACA	64
QY	501	gaatactccaaaatctcatcaacagcgccgtctgtatcgaccagaacacgactccaatct	560
Db	65	CACAACTTCTTCACACTCATCAATCACTTCTCCAAACAACAACAACAACAATCTTCCAGCTA	124
QY	561	gggtaaactlccacgcttcttaataatacatatgttcaaacctggacggttgttcgtgaactta	620
Db	125	CATCAACTTCTTCACACTCATCAATCACTTCTCCAGGTACATCAACTTCTCCAGGTACAAACA	184
QY	621	ccgcatactcgtacataacttaactaatctgttcgacaagaagactgaaagaaagaat	680
Db	185	CTTCTCCAGCTCATCAACTTCTCCAAACAACAACAACAACAACAATCTTCCAAACAACAACA	244
QY	681	caaaagacctgttgcagaacaaccagtlccaatctcgttactcgttgaagaccttcctgggtacta	740
Db	245	CACAAACTTCTTCACACTCATCAACTTCTCCAAACAACAACAACAACAACAATCTTCCAGCTA	304
QY	741	ccctgcagtaagcagaacacacgactactacatgctgaatctgtatcgagaacaaatacgttga	800
Db	305	CATCAACTTCTTCACACTCATCAACAACAACAATCTTCTCCAGGTACATCAACAACAACAACAATCTTCCAG	364
QY	801	cgtaacaaa	809
Db	365	CTACAAACA	373

RESULT	2	
CNS00100		
LOCUS		
DEFINITION		1101 bp DNA linear GSS 03-JUN-1999
ACCESSION		CNS00100
VERSION		Drosophila melanogaster genome survey sequence TET3 end of BAC:
KEYWORDS		BARC32D23 of RPc1-98 library from Drosophila melanogaster (fruit
SOURCE		fly), genomic survey sequence.
ORGANISM		AT068607
		AL068607.1 GI:4958689
		GSS.
		fruit fly.
		Drosophila melanogaster
		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE		1 (bases 1 to 1101)
AUTHORS		Genoscope.
TITLE		Direct Submision
JOURNAL		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage
		BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
COMMENT		- Web : www.genoscope.cns.fr)
		Determination of this BAC-end sequence was carried out as part of a
		collaboration with the Berkeley Drosophila Genome Project (BDGP).
		The BDGP is constructing a physical map of the Drosophila
		melanogaster genome using these BACs. For further information
		please see http://www.fruitfly.org The BDGP Drosophila
		melanogaster BAC library was prepared by Kazuhiro Osoegawa and
		Aaron Mammoser in Pieter de Jong's laboratory in the Department of
		Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
		NY. The library is named RPc1-98 and was constructed by partial
		EcoRI digestion of Drosophila DNA provided by the BDGP from the
		isogenic strain y2: cn bw sp, the same strain used for the BDGP's
		p1 and Est libraries. A more detailed description of the library
		and how to order individual BAC clones, the entire library, or

FEATURES	SOURCE
filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	Location/Qualifiers
1. 1101	

[illegible]

Query Match	4.2%	Score 56.2	DB 12	Length 1101
Best Local Similarity	17.8%	Pred. No. 0.0016		
Matches 101	Conservative 220	Mismatches 245	Indels 3	Gaps 1

[illegible]

RESULT	3
LOCUS	BM277996
DEFINITION	BM277996 537 bp mRNA linear EST 20-DEC-2001
ACCESSION	As_tgz_51F02_SKPI Ascaris suum adult male testis germinal zone from
VERSION	Alan Scott Ascaris suum cDNA clone As_tgz_51F02 5', mRNA sequence.
KEYWORDS	BM277996 GI:17971254
SOURCE	EST.
ORGANISM	pig roundworm.
	Ascaris suum
	Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
REFERENCE	1 (bases 1 to 537)
AUTHORS	Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Gulliano, D., Hall
	, N., Quayle, M. and Barrell, B.


```

Db      342 ACAGTACTTCAGCACTTCACAGCACTTCA 377
RESULT  5
LOCUS   AU060224
DEFINITION AU060224 Dictyostellium discoideum SL (H.Urushihara) Dictyostellium
            discoideum cDNA clone SLA610, mRNA sequence.
ACCESSION AU060224
VERSION   AU060224.1 GI:4881328
KEYWORDS EST.
SOURCE   Dictyostellium discoideum.
ORGANISM Dictyostellium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
AUTHORS  1 (bases 1 to 394)
          Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
          Yoshino,R., Maitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
          Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
          Unpublished (1998)
          Developmental cDNA in Dictyostellium discoideum
          Contact: Hideko Urushihara
          Institute of Biological Sciences
          University of Tsukuba
          3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan
          Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT  = Dictyostellium discoideum cDNA project in Japan.
FEATURES
    source
        1..394
            /organism="Dictyostellium discoideum"
            /strain="AX4"
            /db_xref="taxon:44689"
            /clone="SLA610"
            /dev_stage="slug"
            /dev_stage="slug"
BASE COUNT  223 a      134 c      8 g      28 t      1 others
ORIGIN
Query Match      4.0%; Score 53.2; DB 9; Length 394;
Best Local Similarity 49.3%; Pred. No. 0.0062;
Matches 139; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

```

```

Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
; Ascarididae; Ascaris.
REFERENCE 1 (bases 1 to 447)
AUTHORS  Blaxter,M.L., Parkinson,J., Whittom,C., Daud,J., Gulliano,D., Hall
          N., Quayle,M. and Barrell,B.
          Edinburgh University/Sanger Centre Nematode EST Project
          Unpublished (2000)
          Contact: Blaxter ML
          Institute of Cell, Animal and Population Biology
          University of Edinburgh
          Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
          3JT, UK.
          TEL: +44 131 650 6760
          FAX: +44 131 670 5450
          Email: mark.blaxter@ed.ac.uk
          The library was prepared by Michelle Lizotte-Waniewski for Alan
          Scott, Johns Hopkins University Medical School, Baltimore.
          Sequencing was performed by the Pathogen Sequencing Unit, Sanger
          Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The
          sequence contained a Polya tail (trimmed)
          PCR Primers
          FORWARD: T3
          BACKWARD: T7PL
          Plate: 44 row: C column: 11
          Seq primer: SKPL
          High quality sequence stop: 447.
          Location/Qualifiers
            1..447
                /organism="Ascaris suum"
                /db_xref="taxon:6253"
                /clone="As_tgz_44C11"
                /clone_lib="Ascaris suum adult male testis germinal zone
                from Alan Scott"
                /sex="Male"
                /dev_stage="Adult"
                /note="Vector: Lambda zap II; Site.1: EcoRI; Site.2: XhoI;
                Library was made from dissected testis germinal zone from
                adult male Ascaris suum collected from abattoirs.
                Constructed by Michelle Lizotte-Waniewski for Alan Scott,
                Johns Hopkins University Medical School, Baltimore, MD."
BASE COUNT  173 a      127 c      49 g      98 t
ORIGIN
Query Match      4.0%; Score 52.6; DB 10; Length 447;
Best Local Similarity 49.8%; Pred. No. 0.0092;
Matches 133; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```

Alan Scott Ascaris suum cDNA clone As-tg2_54C05 5', mRNA sequence.
 BM278181
 VERSION BM278181.1 GI:17971439
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.
 1 (bases 1 to 597)
 Blaxter, M.L., Parkinson, J., Whittom, C., Daub, J., Gilliano, D., Hall
 N., Quayle, M., and Barrell, B.
 Edinburgh University/Sanger Centre Nematode EST Project
 Unpublished (2000)
 JOURNAL Contact: Blaxter ML
 COMMENT Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3J7, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared by Michelle Lizotte-Waniewski for Alan
 Scott, Johns Hopkins University Medical School, Baltimore.
 Sequencing was performed by the Pathogen Sequencing Unit, Sanger
 Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).
 PCR Primers
 FORWARD: T3
 BACKWARD: T7PL
 Plate: 54 row: C column: 05
 Seq primer: SKPL
 High quality sequence stop: 502.
 Location/Qualifiers
 1..597
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone="As-tg2_54C05"
 /clone_1lb="Ascaris suum adult male testis germinal zone
 from Alan Scott"
 /sex="Male"
 /dev_stage="Adult"
 /note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
 Library was made from dissected testis germinal zone from
 adult male Ascaris suum collected from abattoirs.
 Constructed by Michelle Lizotte-Waniewski for Alan Scott,
 Johns Hopkins University Medical School, Baltimore, MD."
 BASE COUNT 222 a 167 c 76 g 132 t
 ORIGIN
 Query Match 3.9%; Score 52; DB 10; Length 597;
 Best Local Similarity 49.3%; Pred. No. 0.015;
 Matches 136; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

RESULT 8
 LOCUS BM278805 500 bp mRNA linear EST 20-DEC-2001
 DEFINITION As-tg2_67F10-SKPL Ascaris suum adult male testis germinal zone from
 Alan Scott Ascaris suum cDNA clone As-tg2_67F10 5', mRNA sequence.
 ACCESSION BM278805
 VERSION BM278805.1 GI:17972063
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.
 1 (bases 1 to 500)
 Blaxter, M.L., Parkinson, J., Whittom, C., Daub, J., Gilliano, D., Hall
 N., Quayle, M., and Barrell, B.
 Edinburgh University/Sanger Centre Nematode EST Project
 Unpublished (2000)
 JOURNAL Contact: Blaxter ML
 COMMENT Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3J7, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared by Michelle Lizotte-Waniewski for Alan
 Scott, Johns Hopkins University Medical School, Baltimore.
 Sequencing was performed by the Pathogen Sequencing Unit, Sanger
 Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The
 sequence contained a Polya tail (trimmed)
 PCR Primers
 FORWARD: T3
 BACKWARD: T7PL
 Plate: 67 row: F column: 10
 Seq primer: SKPL
 High quality sequence stop: 500.
 Location/Qualifiers
 1..500
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone="As-tg2_67F10"
 /clone_1lb="Ascaris suum adult male testis germinal zone
 from Alan Scott"
 /sex="Male"
 /dev_stage="Adult"
 /note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
 Library was made from dissected testis germinal zone from
 adult male Ascaris suum collected from abattoirs.
 Constructed by Michelle Lizotte-Waniewski for Alan Scott,
 Johns Hopkins University Medical School, Baltimore, MD."
 BASE COUNT 193 a 144 c 58 g 105 t
 ORIGIN
 Query Match 3.9%; Score 51.8; DB 10; Length 500;
 Best Local Similarity 50.8%; Pred. No. 0.015;
 Matches 153; Conservative 0; Mismatches 142; Indels 6; Gaps 1;

Db 244 TCAGCAATCTCAACATTAACCTTACACACCACTTTCACACAGTAACT 303
 QY 671 aaaaagaatcaagaacctgtacgaacacagcagtccttgatctcgaagactct 730
 Db 304 TCACAACTTCAACAGCACTTTCACCACTTTCACCAAGTAACTTCAACACAGCA 363
 QY 731 g 731
 Db 364 G 364

RESULT 9
 BM004465 511 bp mRNA linear EST 25-OCT-2001
 LOCUS TGE8TZY86f03.y1 TGVEG partially spornulated oocyst cDNA Toxoplasma
 DEFINITION gondii cDNA clone TGE8TZY86f03.y1 5', mRNA sequence.
 ACCESSION BM004465
 VERSION BM004465.1 GI:16439239
 KEYWORDS EST.
 SOURCE Toxoplasma gondii.
 ORGANISM Toxoplasma gondii.
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Sarcocystidae; Toxoplasma.

REFERENCE 1 (bases 1 to 511)
 AUTHORS Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
 Clifton,S., Page,D., Martin,J., Wylie,T., Dante,M., Matra,M.,
 Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Rlter
 ,E., Bennett,J., Franklin,C., Tsagarisvilli,R., Ronko,I., Kennedy
 ,S., Maguire,L., Waterston,R. and Wilson,R.
 Toxoplasma EST project
 Toxoplasma EST project
 Contact: Clifton, S.
 Toxoplasma EST project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: toxowatson.wustl.edu
 Contact David Sibley (toxoeastborcim.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 High quality sequence stop: 433.
 Location/Qualifiers
 1..511
 /organism="Toxoplasma gondii"
 /strain="VEG (Type III)"
 /db_xref="taxon:5811"
 /clone="TGE8TZY86f03.y1"
 /dev_stage="Partially sporulated oocysts"
 /note="vector: Modified plusescript (pBSSK+); Site_1:
 BamHI; Site_2: EcoRI; PolyA mRNA from partially-sporulated
 oocysts was converted to cDNA using the
 template-switching PCR method (SMART cDNA, Clontech Inc.,)
 and sized selected on SizeSep 400 columns (Amersham
 Pharmacia Biotech Inc.). First strand was reverse
 transcribed using the CDS III-oligo-dT primer and a 5'
 template switch primer (Smart IV primer). The product of
 the first strand synthesis was PCR amplified using the
 same primer set and the fragments were digested with SfiI.
 The fragments were size selected, ligated into a modified
 plusescript vector containing directional SfiI sites, and
 electroporated into DH10B or DH12S cells. Vector: SfiI
 sites were added to the multiple cloning region of
 plusescript SK+ between the BamHI/EcoRI sites. The
 modified polylinker has the following sequence:
 5'-GAATTCGGCATACGGCC(G)-n--insert--
 GGCCGCTGGCCACGATCC3'where n=3-4 G nucleotides. Library
 Source: Michael White, Maria E. Jerome,Emily A.Johnson,
 Jay A. Radke, Montana State University. Clone
 Availability: David Sibley, Washington University"

BASE COUNT 213 a 181 c 61 g 56 t

ORIGIN

Query Match 3.8%; Score 50.6; Db 10; Length 511;
 Best Local Similarity 46.9%; Pred. No. 0.031;
 Matches 158; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 426 caaactctcagatgatcaacatctctgactacatacgcgtgactctgctgtaecat 485
 Db 169 CTACTACTACACTACACACCAACCAACCAACCAACCAACCAACCAACCAACCA 228
 QY 486 caccacaatcgtctgataatacctcaaatcattacatacgaagccgctgtgacccagaa 545
 Db 229 CAACGACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 288
 QY 546 accgatctcaatctggtgatacaccacgcttctaataacatcatgltcaaatgacg 605
 Db 289 CAACACACACGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 348
 QY 606 ttgtcgtgaactcaccgctacatctgtgataacatctcaatctgtgtgacaaagact 665
 Db 349 CTACCACACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 408
 QY 666 gaacgaagaagaatcaagaacctgtacgaacacgctccaatctgtatccctgaaga 725
 Db 409 CACACACACGACGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 468
 QY 726 ctctcgtggtgactaccctgacgtacgacgaacacgctac 762
 Db 469 CTACGACACCATCTACGACGACGACGACGACGACGACGACGACGACGACGAC 505

RESULT 10
 BM278657 569 bp mRNA linear EST 20-DEC-2001
 LOCUS As_tgz_65E06_SKPL Ascaris suum adult male testis germinal zone from
 DEFINITION Alan Scott Ascaris suum cDNA clone As_tgz_65E06 5', mRNA sequence.
 ACCESSION BM278657
 VERSION BM278657.1 GI:17971915
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.

REFERENCE 1 (bases 1 to 569)
 AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guillian,D., Hall
 ,N., Quayle,M. and Barrell,B.
 Edinburgh University/Sanger Centre Nematode EST Project
 Unpublished (2000)
 Contact: Blaxter ML
 Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3JT, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared by Michelle Lizotte-Waniewski for Alan
 Scott, Johns Hopkins University Medical School, Baltimore.
 Sequencing was performed by the Pathogen Sequencing Unit, Sanger
 Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The
 sequence contained a PolyA tail (trimmed)
 PCR Primers
 FORWARD: T3
 BACKWARD: 77PL
 Plate: 65 row: E column: 06
 Seq primer: SKPL
 High quality sequence stop: 518.
 Location/Qualifiers
 1..569
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone="As_tgz_65E06"
 /clone_lib="Ascaris suum adult male testis germinal zone
 from Alan Scott"

FEATURES
 source


```

/sex="Male"
/dev.stage="Adult"
/notice=Vector: lambda zap ii: Site.1: EcoRI: Site.2: XhoI
library was made from dissected testis germinal zone from
adult male Ascaris sum collected from abattoirs.
Constructed by Michelle Licotte-Waniewski for Alan Scott,
Johns Hopkins University Medical School, Baltimore, MD."
BASE COUNT      213 a      164 c      69 g      123 t
BRIGIN

```

Query Match	3.8%;	Score 50.4;	DB 10;	Length 569;
Best Local Similarity	48.9%;	Pred. No. 0.037;		
Matches 135; Conservative	0;	Mismatches 141;	Indels 0;	Gaps 0;

QY	431	acttcgagatgaataacatctctgactgaacaaatcgctgagcttcgtttacatcacca	490
Db	119	ACAGTAACCTTCACACACAGCAATTTTCACAGCAACCTCGACAACTTCAATTAACAACTT	178
QY	491	acaaatcgctcgtgaataactccaanaatctacatacaagcgctcgtatcgagccagaaacgga	550
Db	179	GCAACTCTCGAACAACCTTCAATTAACAACTCGACAACTCTTGAGACAACCTCAGCTAC	238
QY	551	ttctcaactctggggaacatccacgctctctataacatcattgtccaacatcggaagcggtgc	610
Db	239	ACTTGAACAACCTTCACAACTCTCATTAACAACTCAATTCATTCACAAACAACTTCAGA	298
QY	611	gtgacactcacccgtatcatctgatactcaatactactcaactctgttcgacaaagaactgaag	670
Db	299	ATTCACAAATTAACCTCAACAACTTGACGACGAACAACCTTCAACAGCAGTAATCTCAACA	358
QY	671	aaaaagaatccaagaccctgtaagacacacagttcca	706
Db	359	ACAGTAACCTTCAGCAACCTTCAACAGCAACCAAGCTTCA	394

RESULT	11
LOCUS	BM181884
DEFINITION	BM181884 641 bp mRNA linear EST 11-DEC-2001 fv2ib11.y1 Sugano SJD adult male Dario rerio cDNA clone 5412044 5' similar to contains element TAR1 repetitive element ;, mRNA

ACCESSION	BM181884	GI:17512842
VERSION	BM181884.1	
KEYWORDS	EST.	

ORGANISM

REFERENCE

Clark, M., Johnson, S.L., Lebrach, H., Lee, R., Li, F., Maria, M., Eddy
1 (bases 1 to 641)
Cypripinae, Danio.
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae.

AUTHORS

'S., Haller, D., Nuccia, J., Martin, J., Beck, C., Wiley, T., Underwood, K., Steptoe, B., Theising, B., Allen, M., Bowers, Y., Patson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

TITLE Washu zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: zdrafishewatson.wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clome
distribution information can be found through the I.M.A.G.E.
Consortium/INLIT, send email to: info@image.jhl.jgov

FEATURES
source

Location/Qualifiers
1. .641

```

/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5412044"
/clone_lib="Sugano SJD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/notes="vector: pME185-FL3; site.1: DraIII (CACCATGCG);
site.2: DraII (CAGCTGTGCG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCTTTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
(FTGTTGGCTCAGTGG), digested and cloned into distinct DraIII
sites of the pME185-FL3 vector (5' site CAGCTGTGCG, 3' site
CACCATGCG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science)
Custom primers for sequencing: 5' end primer
CTTGTGCTCTTAAAGCTGCG and 3' end primer
CGACCTGCACTGCGAGCA."

```

Query Match	3.8%;	Score 50;	DB 10;	Length 641;
Best Local Similarity	46.3%;	Pred. No. 0.049;		
Matches 201; Conservative	0;	Mismatches 230;	Indels 3;	Gaps 1;

OY	229	taaacctatgtagcgaanaacttccaccccttclygatccgtatccgaaatacttc	286
Dδ	175	TACATCAACTTCTCCACAACAACAACACTTCTCCAAACAACAACACTTCTCCAGGTATATC	233
OY	289	aactcatctctctcgacaatgaatgaaccatcaactcaatgcattgataagaacaa---ttct	344
Dδ	235	AACGTCGAACAACAACAACAATTCTCCAGGTACTCATCACTTCAACAACAACAACAATTCTCC	294

Oy 346 ggttggaaagatctctcgtgaactacggtgaaatcatctggaactctgcaggaactcaggaa 405
 Db 295 AGGTACAACCTTCCAGCAGCAACATCACTTCTTCGAGCTACATCAACTTCTCCAAACAC 354

Db 355 AACAACTTCTCGAGCTACATCAACTTCTCCACACAACAACAACCTTCTCGAGCTACATCAAC 414

Db 415 TTCTCAACAACAACAACCTTCTCCAGGTACATCAACTTCAACAACAACAACAACACTTCTCC 472
 Oy 526 ggcgcgtctgatacgacagaaacgcgatactccaaattggtlaacatccacgcgtcttaataac 583
 475 AGCTACATCACTTCAACCAACAACAACACTTCTCCAGGTACATCAACTTCTCCAGCAACAACAATC 530
 476

Db 475 AGCTACATCAACTTCAACACACAACTTCTCCAGCTACAACTTCTCCAGCAACACATC 534

DQ 586 atcattgtcaaacctgcgacgggtgtcgtagaacatccagcgtcacatctgatcaaatacttc 645
| | | | | | | | | | | | | | |
DB 535 AACTTCTCCAGCTACATCACTTAACAACAACAACATTCTTCAGCTACATCAACTTC 594

QY	646	aatcgttcgacaa	659
Db	595	AACACACACACAA	608
RESULT	12		
AI389106			

LOCUS	AI389106	645 bp	mRNA	linear	EST 23-APR-2001
DEFINITION	GH20192.5:prine GH Drosophila melanogaster head por2 melanogaster cDNA clone GH20192_5, mRNA sequence.				
ACCESSION	AI389106				
VERSION	AI389106.2	GI:13758704			

ORGANISM *Drosophila melanogaster*
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Pterygota: Neoptera: Endopterygota: Diptera: Brachyera:
Muscomorpha: Ephydroidea: Drosophilidae, Drosophila:
1 (bases 1 to 645)
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HIMI Drosophila EST Project
Unpublished (2001)
On Jan 28, 1999 this sequence version replaced gi:4203117.

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
Plate: GH.201 row: H column: 8
High quality sequence stop: 627
SOLYA-NO.

FEATURES

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH20192"
/clone_11b="GH Drosophila melanogaster head pOr2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/note="Organ: head; Vector: pOr2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOr2. Plasmid cDNA library."
BASE COUNT      239 a      232 c      94 g      80 t
ORIGIN

```

Query Match	3.7%	Score	49.6	DB	9	Length	645
Best Local Similarity	47.0%	Pred. No.	0.052				
Matches	154	Conservative	0	Mismatches	174	Indels	0
						Gaps	0

Qy	482	ccatcaccaaaacatcgcgttgatbaacccaacatctcatcaaaagcgccctgtatgacc	54111
Db	209	ccaccattggtgttaggaamaaccgacacaaacaaacggtttccctctactaccacaaacacaca	266
Qy	542	agaaacccgatctccatctcgttgatcatccacgcttcctaataacatgttcaaatctg	601
Db	269	acgtgtttccccaactaccacacacacacaaacaaactccacacaaacaaacacccacaca	328
Qy	602	acggctgtcgtgcacatccacgcgtctacatctgcatcaatctactcaatctgttcgacaag	661
Db	329	actacgaacaaacaaacaaactaccacacacaaactacgaacaaacaaacgaatgtttcccgacacaca	388
Qy	662	aactgaaacgaaaaagaatctcaagaacctgtacgaaacccagttccaatctgtgtatcctga	721
Db	389	accacacccacaaacaaacaaacacacacacacacgacgacaaacaaactatctacgacaaacacg	448
Qy	722	aagacatctctggtgtgactccttgtagtcaagtaacaaacgctactactatgctgaatctgacg	781
Db	449	acgacgaacaaacaaacaaacacacacacacacacacacacacacacacacacacacacacac	508
Qy	782	atccgaacaaatactcgttgagctgaacaa	809
Db	509	acaagaactactactacgacaacacacaca	536

RESULT	13
CNS02156	
LOCUS	827 bp DNA linear
DEFINITION	Tetradon nigroviridis genome survey sequence PUC-ORI end of clone 224F10 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION	AL176451
VERSION	AL176451.1 GI:7814508
KEYWORDS	GSS; genome survey sequence, Tetradon nigroviridis.
SOURCE	

ORGANISM *Tetradodon nigroviridis*

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE AUTHORS	TITLE
1 (bases 1 to 827) Roest-Crollius, H., Jalllon, O., Desilva, C., Fitzames, C., Fisher, C., Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish <i>Tetraodon nigroviridis</i>

REFERENCE	2 (bases 1 to 827)
AUTHORS	Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Winkler,P., Brottier,P., Quetier,F., Saulin,W. and Weissenbach,J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 827)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetradon .

FEATURES	source	location/Qualifiers
1. .827		
		/organism="Tetrahodon nigroviridis"
		/db_xref="taxon:99863"
		/clone="224F10"
		/clone_11b="G"
		/note="Genoscope sequence ID : C0AC224DC05SP1-end : PUC-Or1"
BASE COUNT	368 a	203 c 106 g 121 t 29 others
ORIGIN		

Query Match	3.7%	Score 49.2	DB 12	Length 827		
Best Local Similarity	43.8%	Pred. No. 0.086	Mismatches 169	Conservative 9	Indels 0	Gaps 5
426	caaatcctcagatgataaacatctctgactacaatcgctggaatctctgatiacat	485				
166	CRGAGACACACACACAGACACACACACACACACATATACATATACAGCATATACACATATACAA	225				
486	caccaacaatcgctctgataaactccaataattacatacaagcgctctgataccgaa	545				
226	CACACACACACACATATACACACACACACATATACAGCATATACACATATACACACACAA	285				
546	accgatctccatactcgggtaacatccacgcttcaatacactgattcaactgagcgg	605				
286	TACACACACACATATACAGCATATACACACATATACAGCATATACACACACAA	345				
606	tctgtctgactacacgcgtacatcttgatacaaatbaacttcaatctgttcgcaagaact	665				
346	CACACATATACACACACATATACAGCATATACAGCATATACAGCATATACACATATACAA	405				
666	gaacgaaaaaagaatcaaaagaccctgtacgaaacgaagccgacatctgtatccgaaaga	725				
406	CATATACACTGTATACACACATATACACACATATACACACATATACAGCATATACACACAA	465				
726	ctctctgggtgactacctcgtcagtagcaacacgtaactacatgtctgaatctgtatcga	785				
466	TGCACGCGATGTCACACACACACACACACACACACACATATACAGCATATACAGCATATAC	525				
786	gaacaaatcgtttgacgtcaacaatg	811				
526	CACACATATACAGCATATACACACATATACAGCATATACAGCATATACAGCATATAC	585				

RESULT 14

BF050030 481 bp mRNA linear EST 16-OCT-2000
 LOCUS As.tgz.21C12.SKPL Ascaris suum adult male testis germinal zone from
 DEFINITION Alan Scott Ascaris suum cDNA clone As.tgz.21C12 5', mRNA sequence.
 ACCESSION BF050030.1 GI:10803926
 VERSION BF050030.1
 KEYWORDS EST.
 SOURCE pig roundworm.
 ORGANISM Ascaris suum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
 ; Ascarididae; Ascaris.
 REFERENCE 1 (bases 1 to 481)
 AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Gulliano,D., Hall
 N., Quayle,M. and Barrett,B.
 TITLE Edinburgh University/Sanger Centre Nematode EST Project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Blaxter ML
 Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3JF, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared by Michelle Lizotte-Waniewski for Alan
 Scott, Johns Hopkins University Medical School, Baltimore.
 Sequencing was performed by Claire Whitton ICAP, Edinburgh
 PCR PRIMERS
 FORWARD: 13
 BACKWARD: 17PL
 Plate: 21 row: C column: 12
 Seq primer: SKPL
 High quality sequence stop: 456.

FEATURES
 source location/Qualifiers
 1..481
 /organism="Ascaris suum"
 /db_xref="taxon:6253"
 /clone_1lb="As.tgz.21C12"
 /clone_1lb="Ascaris suum adult male testis germinal zone
 from Alan Scott"
 /sex="Male"
 /dev_stage="Adult"
 /note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
 library was made from dissected testis germinal zone from
 adult male Ascaris suum collected from abattoirs.
 Constructed by Michelle Lizotte-Waniewski for Alan Scott,
 Johns Hopkins University Medical School, Baltimore, MD."

BASE COUNT 188 a 149 c 49 g 95 t
 ORIGIN

Query Match 3.7%; Score 49; DB 10; Length 481;
 Best Local Similarity 49.1%; Pred. No. 0.079; Mismatches 135; Indels 0; Gaps 0;

Matches 130; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
 442 atcaatctctgactacatcaatcgatctgactctgactacatcaatcgatctg 501
 104 AACAAAGCAATTTCAACGACACCTCAGCAATTCAATCAACACTTACCACTCTG 163
 502 aataatccaanaattacatcaacgagcgctgctgacgacgaagaacgatctcaatctg 561
 164 AACAACTTCAATTAACACTCAGCAACTCTGAGCAACTCTGAGCAACTCTGAGCAAC 223
 562 ggttaaatccagctcttaatacatcatgcttaaatctgagaggttgcgtgacacac 621
 224 CTCACACACCTCAATTAACACCTCAATCTCAACAACTCTCAACAACTCTCAACAA 283
 622 cgtataatctgataatcaatctgctgacacgaagaactgaagaagaagaatc 681
 284 AACCTCAACACCTTGAACGACGACCACTTCAACAGCAGTAACTTCAACAGCACTTCA 343
 682 aaagacctgtagacacacagctca 706

Db 344 AGCAACTTCAACAGCAACAACTTCA 368
 RESULT 15
 LOCUS AU060230 628 bp mRNA linear EST 20-MAY-1999
 DEFINITION AU060230 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
 discoideum cDNA clone SLA628, mRNA sequence.
 ACCESSION AU060230
 VERSION AU060230.1 GI:4881334
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum.
 Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
 REFERENCE 1 (bases 1 to 628)
 AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
 Yoshino,R., Mita,B.N., Pl.M., Saito,T., Takemoto,K., Yasukawa,H.,
 Williams,J., Maeda,M., Takeuchi,I., Ochial,H. and Tanaka,Y.
 TITLE Developmental cDNA in Dictyostelium discoideum
 JOURNAL Unpublished (1998)
 COMMENT Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
 Email: d402huesakura.cc.tsukuba.ac.jp
 PROJECT - Dictyostelium discoideum cDNA project in Japan.

FEATURES
 source location/Qualifiers
 1..628
 /organism="Dictyostelium discoideum"
 /strain="Ax4"
 /db_xref="taxon:44689"
 /clone_1lb="SLA628"
 /dev_stage="slug"
 BASE COUNT 270 a 160 c 56 g 142 t
 ORIGIN

Query Match 3.7%; Score 48.8; DB 9; Length 628;
 Best Local Similarity 48.6%; Pred. No. 0.098; Mismatches 142; Indels 0; Gaps 0;
 Matches 134; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

426 caaatactcagatgatacatctgactacatcaatcgatctgactacat 485
 78 CAACCCACATCCACCAACAAACAAACAAACAAACAAACAAACAAACAA 137
 486 caccacaatcgctgataactcaaatctacatcaacgagcgctgactgacagaa 545
 138 CAATTAACAAATTAACAAACAAACAAACAAACAAACAAACAAACAA 197
 546 accgatctcaatctggtgataatccagctcttaataatcatatgttcaactgagcgg 605
 198 CAACAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCA 257
 606 ttgctgacacatcaacgctacatctgatacaataatctgctgacaaagact 665
 258 ATTCAACAAATATATCAACAAACAAACAAACAAACAAACAAATCAAC 317
 666 gaacgaanaaagaatcaagaacctgtagacgaacca 701
 318 CAACAAACAAATATCAAAATGTTATTCACAAACAA 353

Search completed: September 16, 2002, 19:28:19
 Job time: 8376 sec

